ANJU PANWAR

MATHEMATICAL ANALYSIS

ANJU PANWAR

MATHEMATICAL ANALYSIS

CONTENTS

UNIT IV: TAYLOR THEOREM

UNIT – I

THE RIEMANN-STIELTJES INTEGRAL

Structure

1.0 Introduction

In this unit, we will deal with the Riemann-Stieltjes integral and study its existence and properties. The Riemann-Stieltjes integral is a generalization of Riemann integral named after Bernhard Riemann and Thomas Joannes Stieltjes. The reason for introducing this concept is to get a more unified approach to the theory of random variables. Fundamental Theorem of the Integral Calculus is discussed later on.

1.1 Unit Objectives

After going through this unit, one will be able to

- define Riemann-Stieltjes integral and characterize its properties.
- recognize Riemann-Stieltjes integral as a limit of sums.
- know about Fundamental Theorem of the Integral Calculus and Mean Value Theorems .
- understand the concept of Rectifiable Curves

1.2 Riemann-Stieltjes integral

We have already studied the Riemann integrals in our undergraduate level studies in Mathematics. Now we consider a more general concept than that of Riemann. This concept is known as Riemann-Stieltjes integral which involve two functions f and α . In what follows, we shall consider only real-valued functions.

1.2.1 Definitions and Notations

Definition1. Let $[a,b]$ be a given interval. By a **partition** (or subdivision) P of $[a,b]$, we mean a finite set of points

$$P = \{x_0, x_1, ..., x_n\}$$

such that

$$a = x_0 \le x_1 \le x_2 \le ....x_{n-1} \le x_n = b .$$

Definition 2. A partition P^* of $[a,b]$ is said to be **finer than P** (or a refinement of P) if $P^* \supseteq P$, that is , if every point of P is a point of P^* i.e. $P \subseteq P^*$.

Definition 3. The P_1 and P_2 be two partitions of an interval $[a,b]$. Then a partition P^* is called their **common refinement** of P_1 and P_2 if $P^* = P_1 \cup P_2$.

Definition 4. The length of the largest subinterval of a partition $P = \{x_0, x_1, ..., x_n\}$ of $[a,b]$ is called the **Norm** (or **Mesh)** of P. We denote norm of P by $|P|$. Thus

$$|P| = \max \Delta x_i = \max\{x_i - x_{i-1} : i = 1, 2, ..., n\}$$

We notice that if $P^* \supseteq P$, then $|P^*| \le |P|$. Thus refinement of a partition decreases its norm.

Definition 5. Lower and Upper Riemann-Stieltjes Sums and Integrals

Let f be a bounded real function defined on a closed interval $[a,b]$. Corresponding to each partition P of $[a,b]$, we put

$$M_i = \text{lub } f(x) \qquad (x_{i-1} \le x \le x_i)$$
$$m_i = \text{glb } f(x) \qquad (x_{i-1} \le x \le x_i).$$

Let α be monotonically increasing function on $[a,b]$. Then α is bounded on $[a,b]$ since $\alpha(a)$ and $\alpha(b)$ are finite.

Corresponding to each partition P={$x_0, x_1, \ldots\ldots\ldots, x_n$} of [a, b], we put

$$\Delta\alpha_i = \alpha(x_i) - \alpha(x_{i-1}).$$

The monotonicity of α implies that $\Delta\alpha_i \geq 0$.

For any real valued bounded function f on $[a,b]$, we take

$$L(P,f,\alpha) = \sum_{i=1}^{n} m_i \Delta\alpha_i$$

$$U(P,f,\alpha) = \sum_{i=1}^{n} M_i \Delta\alpha_i \, ,$$

where m_i and M_i are bounds of f defined above. The sums $L(P,f,\alpha)$ and $U(P,f,\alpha)$ are respectively called **Lower Stieltjes sum** and **Upper Stieltjes sum** corresponding to the partition P. We further define

$$\underline{\int_a^b} f d\alpha = \text{lub}\, L(P,f,\alpha)$$

$$\overline{\int_a^b} f d\alpha = \text{glb}\, U(P,f,\alpha),$$

where lub and glb are taken over all possible partitions P of $[a,b]$. Then $\underline{\int_a^b} f d\alpha$ and $\overline{\int_a^b} f d\alpha$ are respectively called Lower integral and Upper integrals of f with respect to α.

If the lower and upper integrals are equal, then their common value, denoted by $\int_a^b f d\alpha$, is called the **Riemann-Stieltjes integral** of f with respect to α, over $[a,b]$ and in that case we say that f is integrable with respect to α, in the Riemann sense and we write $f \in \Re(\alpha)$.

The functions f and α are known as the **integrand** and the **integrator** respectively.

In the special case, when $\alpha(x) = x$, the Riemann-Stieltjes integral reduces to Riemann-integral. In such a case, we write $L(P,f), U(P,f), \underline{\int_a^b} f, \overline{\int_a^b} f$ and $f \in \Re$ respectively in place of $L(P,f,\alpha), U(P,f,\alpha)$, $\underline{\int_a^b} f d\alpha, \overline{\int_a^b} f d\alpha$ and $f \in \Re(\alpha)$.

Clearly, the numerical value of $\int f d\alpha$ depends only on f, α, a and b and does not depend on the symbol x. In fact x is a "dummy variable" and may be replaced by any other convenient symbol.

1.3 Existence and properties

1.3.1 In this section, we shall study characterization of upper and lower Stieltjes sums, and upper and lower Stieltjes integrals.

The next theorem shows that for increasing function α, refinement of the partition increases the lower sums and decreases the upper sums.

Theorem1. If P^* is a refinement of P, f is bounded real valued function on $[a,b]$ and α is monotonically increasing function defined on $[a,b]$. Then,

$$L(P^*, f, \alpha) \geq L(P, f, \alpha)$$
$$U(P^*, f, \alpha) \leq U(P, f, \alpha).$$

Proof. Let $P = \{x_0, x_1, ..., x_n\}$ be a partition of $[a,b]$. Further, let P^* be a refinement of P having one more point.

Let x^* be such that point in the sub-interval $[x_{i-1}, x_i]$ that is $P^* = \{x_0, x_1, ..., x_{i-1}, x^*, x_i, ..., x_n\}$. Then, let

$$m_i = \text{g.l.b. of f in } [x_{i-1}, x_i]$$

$$w_1 = \text{g.l.b. of f in } \left[x_{i-1}, x^*\right]$$

$$w_2 = \text{g.l.b of f in } \left[x^*, x_i\right].$$

Obviously, $m_i \leq w_1$; $m_i \leq w_2$.

Then,

$$L(P, f, \alpha) = m_1 \Delta\alpha_1 + m_2 \Delta\alpha_2 + ... + m_{i-1}\Delta\alpha_{i-1} + m_i \Delta\alpha_i + ... + m_n \Delta\alpha_n$$

$$L(P^*, f, \alpha) = m_1 \Delta\alpha_1 + m_2 \Delta\alpha_2 + ... + m_{i-1}\Delta\alpha_{i-1} + w_1\left[\alpha(x^*) - \alpha(x_{i-1})\right] + w_2\left[\alpha(x_i) - \alpha(x^*)\right] +$$
$$m_{i+1}\Delta\alpha_{i+1} + ... + m_n \Delta\alpha_n$$

Thus,

$$L(P^*, f, \alpha) - L(P, f, \alpha) = w_1(\alpha(x^*) - \alpha(x_{i-1})) + w_2(\alpha(x_i) - \alpha(x^*)) - m_i(\alpha(x_i) - \alpha(x_{i-1}))$$
$$= (w_1 - m_i)(\alpha(x^*) - \alpha(x_{i-1})) + (w_2 - m_i)(\alpha(x_i) - \alpha(x^*))$$

Now, $w_1 - m_i \geq 0$; $w_2 - m_i \geq 0$

Also, α is monotonically increasing function and $x_{i-1} \le x^* \le x_i$. So,

$$\alpha(x^*) - \alpha(x_{i-1}) \ge 0$$
$$\alpha(x_i) - \alpha(x^*) \ge 0$$

$$\Rightarrow L(P^*, f, \alpha) - L(P, f, \alpha) \ge 0$$
$$\Rightarrow L(P^*, f, \alpha) \ge L(P, f, \alpha)$$

Similarly, $U(P^*, f, \alpha) \le U(P, f, \alpha)$.

If P^* contains more points, then similar process holds and so the result follows.

Theorem 2. For any two partitions P_1 and P_2 of $[a,b]$, let f be a bounded real valued function defined on $[a,b]$ and α is monotonically increasing function defined on $[a,b]$, then

$$L(P_1, f, \alpha) \le U(P_2, f, \alpha).$$

Proof. Let P be the common refinement of P_1 and P_2, that is, $P = P_1 \cup P_2$. Then, using Theorem 1, we have

$$L(P_1, f, \alpha) \le L(P, f, \alpha) \le U(P, f, \alpha) \le U(P_2, f, \alpha).$$

Remark 1. If $m \le f(x) \le M$. Then,

$$m(\alpha(b) - \alpha(a)) \le L(P, f, \alpha) \le U(P, f, \alpha) \le M(\alpha(b) - \alpha(a)).$$

Proof . By hypothesis

$$m \le m_i \le M_i \le M$$

$$\Rightarrow m\Delta\alpha_i \le m_i\Delta\alpha_i \le M_i\Delta\alpha_i \le M\Delta\alpha_i$$

$$\Rightarrow m\sum_{i=1}^{n} \Delta\alpha_i \le m_i\sum_{i=1}^{n} \Delta\alpha_i \le M_i\sum_{i=1}^{n} \Delta\alpha_i \le M\sum_{i=1}^{n} \Delta\alpha_i$$

$$\Rightarrow m(\alpha(b) - \alpha(a)) \le L(P, f, \alpha) \le U(P, f, \alpha) \le M(\alpha(b) - \alpha(a)).$$

Theorem 3. If f is bounded real valued function defined on $[a,b]$ and α is monotonic function defined on $[a,b]$. Then,

$$\underline{\int_a^b} f d\alpha \le \overline{\int_a^b} f d\alpha.$$

Proof . Let $P[a,b]$ denotes the set of all partition of $[a,b]$. For $P_1, P_2 \in P[a,b]$, we know that

$$L(P_1,f,\alpha)\le U(P_2,f,\alpha) \qquad (1)$$

This holds for each $P_1 \in P[a,b]$, keeping P_2 fixed, it follows from (1) that $U(P_2,f,\alpha)$ is an upper bound of the set $\{L(P_1,f,\alpha): P_1 \in P[a,b]\}$.

But least upper bound of this set is $\underline{\int_a^b} f(x)d\alpha$.

i.e, $$\underline{\int_a^b} f(x)d\alpha = l.u.b.\{L(P_1,f,\alpha): P_1 \in P[a,b]\}$$

Since, least upper bound $\le$ any upper bound

$$\underline{\int_a^b} f(x)\,d\alpha \le U(P_1,f,\alpha) \qquad (2)$$

This holds for each $P_2 \in P[a,b]$. So, it follows from (2) that $\underline{\int_a^b} f(x)d\alpha$ is a lower bound of the set $\{U(P_2,f,\alpha): P_2 \in P[a,b]\}$.

But greatest lower bound of this set is $\overline{\int_a^b} fd\alpha$.

i.e, $$\overline{\int_a^b} fd\alpha = g.l.b.\{U(P_2,f,\alpha): P_2 \in P[a,b]\}$$

Since, any lower bound $\le$ greatest lower bound.

So, $$\underline{\int_a^b} fd\alpha \le \overline{\int_a^b} fd\alpha.$$

Example 1. Let $\alpha(x)=x$ and define f on $[0,1]$ by

$$f(x)=\begin{cases}1, x\in Q\\ 0, x\notin Q\end{cases}$$

Then for every partition P of $[0,1]$, we have

$m_i=0$, $M_i=1$, because every subinterval $[x_{i-1},x_i]$ contain both rational and irrational number. Therefore

$$L(P,f,\alpha)=\sum_{i=1}^{n} m_i \Delta x_i$$

$$=0$$

$$U(P,f,\alpha)=\sum_{i=1}^{n} M_i \Delta x_i$$

$$=\sum_{i=1}^{n}(x_i-x_{i-1})=x_n-x_0=1-0=1$$

Hence, in this case

$$\underline{\int} f d\alpha \le \overline{\int} f d\alpha .$$

Theorem 4. Let α is monotonically increasing on $[a,b]$ then $f\in\Re(\alpha)$ iff for any $\in>0$, there exists a partition P of $[a,b]$ such that

$$U(P,f,\alpha)-L(P,f,\alpha)<\varepsilon$$

Proof. The condition is necessary:

Let f be integrable on $[a,b]$ i.e, $f\in\Re(\alpha)$ on $[a,b]$,

$$\text{so that } \underline{\int_a^b} f d\alpha=\overline{\int_a^b} f d\alpha=\int_a^b f d\alpha \qquad (1)$$

Let $\varepsilon>0$ be given.

Since $\underline{\int_a^b} f d\alpha=\sup\{L(P,f,\alpha): P$ is a partition of [a,b]}

So, by definition of l.u.b., there exists a partition P_1 of $[a,b]$ such that

$$L(P_1,f,\alpha)>\underline{\int_a^b} f d\alpha-\frac{\varepsilon}{2}=\int_a^b f d\alpha-\frac{\varepsilon}{2} \qquad \text{(by (1))}$$

$$\Rightarrow L(P_1,f,\alpha)>\int_a^b f d\alpha-\frac{\varepsilon}{2}$$

$$\Rightarrow L(P_1,f,\alpha)+\frac{\varepsilon}{2}>\int_a^b f d\alpha \qquad (2)$$

Again,

Since $\bar{\int_a^b} f d\alpha = \inf\{U(P,f,\alpha) : P$ is a partition of [a,b]}.

By the definition of g.l.b., there exists a partition P_2 of $[a,b]$ such that

$$U(P_2,f,\alpha) < \bar{\int_a^b} f d\alpha + \frac{\varepsilon}{2} = \int_a^b f d\alpha + \frac{\varepsilon}{2}$$

$$\Rightarrow U(P_2,f,\alpha) < \int_a^b f d\alpha + \frac{\varepsilon}{2} \qquad (3)$$

Let $P = P_1 \cup P_2$ be the common refinement of P_1 and P_2 , so that

$$U(P,f,\alpha) \le U(P_2,f,\alpha) \qquad (4)$$

And $$L(P_1,f,\alpha) \le L(P,f,\alpha) \qquad (5)$$

Now, we have

$$U(P,f,\alpha) \le U(P_2,f,\alpha) < \int_a^b f d\alpha + \frac{\varepsilon}{2} \qquad \text{(by (3))}$$

$$< L(P_1,f,\alpha) + \frac{\varepsilon}{2} + \frac{\varepsilon}{2} \qquad \text{(by (2))}$$

$$= L(P_1,f,\alpha) + \varepsilon$$

$$\le L(P,f,\alpha) + \varepsilon \qquad \text{(by (5))}$$

$$\Rightarrow U(P,f,\alpha) < L(P,f,\alpha) + \varepsilon$$

or $U(P,f,\alpha) - L(P,f,\alpha) < \varepsilon$.

The condition is sufficient:

Let $\varepsilon > 0$ be any number. Let P be a partition of $[a,b]$ such that

$$U(P,f,\alpha) - L(P,f,\alpha) < \varepsilon \qquad (6)$$

Since lower integral condition exceed the upper integral.

So, $\underline{\int_a^b} f d\alpha \le \bar{\int_a^b} f d\alpha$.

$$\Rightarrow \overline{\int_a^b} f d\alpha - \underline{\int_a^b} f d\alpha \geq 0 \qquad (7)$$

Now, we know that

$$L(P,f,\alpha) \leq \underline{\int_a^b} f d\alpha \leq \overline{\int_a^b} f d\alpha \leq U(P,f,\alpha)$$

$$\Rightarrow \overline{\int_a^b} f d\alpha - \underline{\int_a^b} f d\alpha \leq U(P,f,\alpha) - L(P,f,\alpha) < \varepsilon \qquad (8)$$

From (7) and (8), we have

$$0 \leq \overline{\int_a^b} f d\alpha - \underline{\int_a^b} f d\alpha < \varepsilon$$

The non – negative number $\overline{\int_a^b} f d\alpha - \underline{\int_a^b} f d\alpha$ being less than every positive number ε must be zero,

i.e, $\overline{\int_a^b} f d\alpha - \underline{\int_a^b} f d\alpha = 0$

$\Rightarrow \overline{\int_a^b} f d\alpha = \underline{\int_a^b} f d\alpha$.

1.3.2 In this section, we shall discuss integrability of continuous and monotonic functions along with properties of Riemann-Stieltjes integrals.

Theorem 1. If f is continuous on $[a,b]$, then

(i) $f \in \Re(\alpha)$

(ii) to every $\in > 0$ there corresponds a $\delta > 0$ such that

$$\left| \sum_{i=1}^{n} f(t_i)\Delta\alpha_i - \int_a^b f d\alpha \right| < \in$$

for every partition P of $[a,b]$ with $|P| < \delta$ and for all $t_i \in [x_{i-1}, x_i]$.

Proof. (i) Let $\in > 0$ and select $\eta > 0$ such that

$$\eta[\alpha(b)-\alpha(a)] < \in \qquad (1)$$

which is possible by monotonicity of α on $[a,b]$. Also f is continuous on compact set $[a,b]$. Hence f is uniformly continuous on $[a,b]$. Therefore there exists a $\delta > 0$ such that

$$|f(x)-f(t)| < \eta \text{ whenever } |x-t| < \delta \text{ for all } x \in [a,b],\ t \in [a,b] \qquad (2).$$

Choose a partition P with $|P| < \delta$. Then (2) implies

$$M_i - m_i \le \eta \qquad (i = 1, 2, \ldots\ldots, n)$$

Hence

$$\begin{aligned} U(P,f,\alpha) - L(P,f,\alpha) &= \sum_{i=1}^{n} M_i \Delta\alpha_i - \sum_{i=1}^{n} m_i \Delta\alpha_i \\ &= \sum_{i=1}^{n} (M_i - m_i)\Delta\alpha_i \le \eta \sum_{i=1}^{n} \Delta\alpha_i \\ &= \eta \sum_{i=1}^{n} [\alpha_i(x_i) - \alpha(x_{i-1})] \\ &= \eta[\alpha(b)-\alpha(a)] \\ &< \eta . \frac{\in}{\eta} = \in, \end{aligned}$$

which is necessary and sufficient condition for $f \in \Re(\alpha)$.

(ii) We have

$$L(P,f,\alpha) \le \sum_{i=1}^{n} f(t_i)\Delta\alpha_i \le U(P,f,\alpha)$$

and

$$L(P,f,\alpha) \le \int_a^b f d\alpha \le U(P,f,\alpha)$$

Since $f \in \Re(\alpha)$, for each $\in > 0$ there exists $\delta > 0$ such that for all partition P with $|P| < \delta$, we have

$$U(P,f,\alpha) - L(P,f,\alpha) < \in$$

Thus

$$\left| \sum_{i=1}^{n} f(t_i)\Delta\alpha_i - \int_a^b f d\alpha \right| < U(P,f,\alpha) - L(P,f,\alpha)$$

$$<\in$$

Thus for continuous functions f, $\lim_{|P|\to 0}\sum_{i=1}^{n} f(t_i)\Delta\alpha_i$ exists and is equal to $\int_a^b f d\alpha$.

Theorem 2. If f is monotonic on $[a,b]$ and if α is both monotonic and continuous on $[a,b]$, then $f \in \Re(\alpha)$.

Proof. Let $\in$ be a given positive number. For any positive integer n, choose a partition P of $[a,b]$ such that

$$\Delta\alpha_i = \frac{\alpha(b)-\alpha(a)}{n} \ (i = 1,2,......,n).$$

This is possible since α is continuous and monotonic on $[a,b]$ and so assumes every value between its bounds $\alpha(a)$ and $\alpha(b)$. It is sufficient to prove the result for monotonically increasing function f, the proof for monotonically decreasing function being analogous. The bounds of f in $[x_{i-1}, x_i]$ are then

$$m_i = f(x_{i-1}),\ M_i = f(x_i),\ i = 1,2,......,n.$$

Hence

$$\begin{aligned} U(P,f,\alpha) - L(P,f,\alpha) &= \sum_{i=1}^{n}(M_i - m_i)\Delta\alpha_i \\ &= \frac{\alpha(b)-\alpha(a)}{n}\sum_{i=1}^{n}(M_i - m_i) \\ &= \frac{\alpha(b)-\alpha(a)}{n}\sum_{i=1}^{n}[f(x_i) - f(x_{i-1})] \\ &= \frac{\alpha(b)-\alpha(a)}{n}[f(b) - f(a)] \\ &< \in \quad \text{for large n.} \end{aligned}$$

Hence $f \in \Re(\alpha)$.

Example 1. Let f be a function defined by

$$f(x^*) = 1 \text{ and } f(x) = 0 \ \text{ for } x \neq x^*,\ a \le x^* \le b.$$

Suppose α is increasing on $[a,b]$ and is continuous at x^*. Then $f \in \Re(\alpha)$ over $[a,b]$ and $\int_a^b f d\alpha = 0$.

Solution. Let $P = \{x_0, x_1, ..., x_n\}$ be a partition of $[a,b]$ and let $x^* \in \Delta x_i$. Since α is continuous at x^*, to each $\in > 0$ there exists $\delta > 0$ such that

$$\left|\alpha(x)-\alpha(x^*)\right|<\frac{\in}{2} \quad \text{whenever} \quad \left|x-x^*\right|<\delta$$

Again since α is an increasing function,

$$\alpha(x)-\alpha(x^*)<\frac{\in}{2} \qquad \text{for} \qquad 0<x-x^*<\delta$$

and

$$\alpha(x^*)-\alpha(x)<\frac{\in}{2} \qquad \text{for} \qquad 0<x-x^*<\delta$$

Then for a partition P of $[a,b]$,

$$\Delta\alpha_i=\alpha(x_i)-\alpha(x_{i-1})$$

$$=\alpha(x_i)-\alpha(x^*)+\alpha(x^*)-\alpha(x_{i-1})$$

$$<\frac{\in}{2}+\frac{\in}{2}=\in.$$

Therefore $\sum_{i=1}^{n} f(t_i)\Delta\alpha_i=\begin{cases}0, t_i\neq x^*\\ \Delta\alpha_i, t_i=x^*\end{cases}$

that is,

$$\left|\sum_{i=1}^{n} f(t_i)\Delta\alpha_i-0\right|<\in$$

Hence

$$\lim_{|P|\to 0}\sum_{i=1}^{n} f(t_i)\Delta\alpha_i=\int_a^b f d\alpha=0$$

and so $f\in\Re(\alpha)$ and $\int_a^b f d\alpha=0$.

Theorem 3. Let $f_1\in\Re(\alpha)$ and $f_2\in\Re(\alpha)$ on $[a,b]$, then $(f_1+f_2)\in\Re(\alpha)$ and

$$\int_a^b (f_1+f_2)d\alpha=\int_a^b f_1 d\alpha+\int_a^b f_2 d\alpha$$

Proof. Let $P=\{a=x_0,x_1,......,x_n=b\}$ be any partition of $[a,b]$. Suppose further that $M_i^{'},m_i^{'},M_i^{''},m_i^{''}$ and M_i,m_i are the bounds of f_1,f_2 and f_1+f_2 respectively in the subinterval $[x_{i-1},x_i]$. If $\alpha_1,\alpha_2\in[x_{i-1},x_i]$, then

$$\left|[f_1(\alpha_2)+f_2(\alpha_2)]-[f_1(\alpha_1)+f_2(\alpha_1)]\right|$$

$$\leq |f_1(\alpha_2) - f_1(\alpha_1)| + |f_2(\alpha_2) - f_2(\alpha_1)|$$

$$\leq (M_i^{'} - m_i^{'}) + (M_i^{''} - m_i^{''})$$

Therefore, since this hold for all $\alpha_1, \alpha_2 \in [x_{i-1}, x_i]$, we have

$$M_i - m_i \leq (M_i^{'} - m_i^{'}) + (M_i^{''} - m_i^{''}) \qquad (1)$$

Since $f_1, f_2 \in \Re(\alpha)$, there exists a partition P_1 and P_2 of $[a,b]$ such that

$$\begin{cases} U(P_1, f_1, \alpha) - L(P_1, f_1, \alpha) < \dfrac{\in}{2} \\ U(P_2, f_2, \alpha) - L(P_2, f_2, \alpha) < \dfrac{\in}{2} \end{cases} \qquad (2)$$

These inequalities hold if P_1 and P_2 are replaced by their common refinement P.

Thus using (1), we have for $f = f_1 + f_2$,

$$\begin{aligned} U(P, f, \alpha) - L(P, f, \alpha) &= \sum_{i=1}^{n} (M_i - m_i)\Delta\alpha_i \\ &\leq \sum_{i=1}^{n} (M_i^{'} - m_i^{'})\Delta\alpha_i + \sum_{i=1}^{n} (M_i^{''} - m_i^{''})\Delta\alpha_i \\ &< \frac{\in}{2} + \frac{\in}{2} \text{ (using (2))} \\ &= \in . \end{aligned}$$

Hence $f = f_1 + f_2 \in \Re(\alpha)$.

Further, we note that

$$m_i^{'} - m_i^{''} \leq m_i \leq M_i \leq M_i^{'} + M_i^{''}$$

Multiplying by $\Delta\alpha_i$ and adding for $i = 1, 2, ..., n,$ we get

$$\begin{aligned} L(P, f_1, \alpha) - L(P, f_2, \alpha) &\leq L(P, f, \alpha) \leq U(P, f, \alpha) \\ &\leq U(P, f_1, \alpha) \leq U(P, f_1, \alpha) + U(P, f_2, \alpha) \end{aligned} \qquad (3)$$

Also

$$U(P, f_1, \alpha) < \int_a^b f_1 d\alpha + \frac{\in}{2} \qquad (4)$$

$$U(P, f_2, \alpha) < \int_a^b f_2 d\alpha + \frac{\in}{2} \qquad (5)$$

Combining (3), (4) and (5), we have

$$\int_a^b f d\alpha \le U(P,f,\alpha) \le U(P,f_1,\alpha) + U(P,f_2,\alpha)$$

$$< \int_a^b f_1 d\alpha + \int_a^b f_2 d\alpha + \frac{\in}{2} + \frac{\in}{2}$$

Since $\in$ is arbitrary positive number, we have

$$\int_a^b f d\alpha \le \int_a^b f_1 d\alpha + \int_a^b f_2 d\alpha \qquad (6)$$

Proceeding with $(-f_1), (-f_2)$ in place of f_1 and f_2 respectively, we have

$$\int_a^b (-f) d\alpha \le \int_a^b (-f_1) d\alpha + \int_a^b (-f_2) d\alpha$$

or

$$\int_a^b f d\alpha \ge \int_a^b f_1 d\alpha + \int_a^b f_2 d\alpha \qquad (7)$$

Now (6) and (7) yield

$$\int_a^b f d\alpha = \int_a^b (f_1 + f_2) d\alpha = \int_a^b f_1 d\alpha + \int_a^b f_2 d\alpha .$$

Theorem 4. If $f \in \Re(\alpha)$ and $f \in \Re(\beta)$ then $f \in \Re(\alpha+\beta)$ and

$$\int_a^b f d(\alpha+\beta) = \int_a^b f d\alpha + \int_a^b f d\beta .$$

Proof. Since $f \in \Re(\alpha)$ and $f \in \Re(\beta)$, there exist partitions P_1 and P_2 such that

$$U(P_1,f,\alpha) - L(P_1,f,\alpha) < \frac{\in}{2}$$

$$U(P_2,f,\beta) - L(P_2,f,\beta) < \frac{\in}{2}$$

These inequalities hold if P_1 and P_2 are replaced by their common refinement P.

Also

$$\Delta(\alpha_i + \beta_i) = [\alpha(x_i) - \alpha(x_{i-1})] + [\beta(x_i) - \beta(x_{i-1})]$$

Hence, if M_i and m_i are bounds of f in $[x_{i-1}, x_i]$,

$$U(P,f,(\alpha+\beta))-L(P,f,(\alpha+\beta))=\sum_{i=1}^{n}(M_i-m_i)\Delta(\alpha_i+\beta_i)$$

$$=\sum_{i=1}^{n}(M_i-m_i)\Delta\alpha_i+\sum_{i=1}^{n}(M_i-m_i)\Delta\beta_i$$

$$<\frac{\in}{2}+\frac{\in}{2}=\in.$$

Hence $f\in\Re(\alpha+\beta)$.

Further

$$U(P,f,\alpha)<\int_a^b fd\alpha+\frac{\in}{2}$$

$$U(P,f,\beta)<\int_a^b fd\beta+\frac{\in}{2}$$

and

$$U(P,f,\alpha+\beta)=\sum M_i\Delta\alpha_i+\sum M_i\Delta\beta_i$$

Also, then

$$\int_a^b fd(\alpha+\beta)\le U(P,f,\alpha+\beta)=U(P,f,\alpha)+U(P,f,\beta)$$

$$<\int_a^b fd\alpha+\frac{\in}{2}+\int_a^b fd\beta+\frac{\in}{2}$$

$$=\int_a^b fd\alpha+\int_a^b fd\beta+\in$$

Since $\in$ is arbitrary positive number, therefore

$$\int_a^b fd(\alpha+\beta)\le\int_a^b fd\alpha+\int_a^b fd\beta.$$

Replacing f by $-f$, this inequality is reversed and hence

$$\int_a^b fd(\alpha+\beta)=\int_a^b fd\alpha+\int_a^b fd\beta.$$

Theorem 5. If $f\in\Re(\alpha)$ on $[a,b]$, then $f\in\Re(\alpha)$ on $[a,c]$ and $f\in\Re(\alpha)$ on $[c,b]$ where c is a point of $[a,b]$ and

$$\int_a^b fd\alpha=\int_a^c fd\alpha+\int_c^b fd\alpha.$$

Proof. Since $f \in \Re(\alpha)$, there exists a partition P such that

$$U(P, f, \alpha) - L(P, f, \alpha) < \in, \quad \in > 0.$$

Let P^* be a refinement of P such that $P^* = P \cup \{c\}$. Then

$$L(P, f, \alpha) \leq L(P^*, f, \alpha) \leq U(P^*, f, \alpha) \leq U(P, f, \alpha)$$

which yields

$$U(P^*, f, \alpha) - L(P^*, f, \alpha) \leq U(P, f, \alpha) - L(P, f, \alpha) \tag{1}$$

$$< \in$$

Let P_1 and P_2 denote the sets of points of P^* between $[a, c]$, $[c, b]$ respectively. Then P_1 and P_2 are partitions of $[a, c]$ and $[c, b]$ and $P^* = P_1 \cup P_2$. Also

$$U(P^*, f, \alpha) = U(P_1, f, \alpha) + U(P_2, f, \alpha) \tag{2}$$

and

$$L(P^*, f, \alpha) = L(P_1, f, \alpha) + L(P_2, f, \alpha) \tag{3}$$

Then (1), (2) and (3) imply that

$$U(P^*, f, \alpha) - L(P^*, f, \alpha) = [U(P_1, f, \alpha) - L(P_1, f, \alpha)] + [U(P_2, f, \alpha) - L(P_2, f, \alpha)]$$

$$< \in$$

Since each of $U(P_1, f, \alpha) - L(P_1, f, \alpha)$ and $U(P_2, f, \alpha) - L(P_2, f, \alpha)$ is non-negative, it follows that

$$U(P_1, f, \alpha) - L(P_1, f, \alpha) < \in$$

and

$$U(P_2, f, \alpha) - L(P_2, f, \alpha) < \in$$

Hence f is integrable on $[a, c]$ and $[c, b]$.

Taking inf for all partitions, the relation (2) yields

$$\int_a^{\bar{b}} f d\alpha \geq \int_a^{\bar{c}} f d\alpha + \int_c^{\bar{b}} f d\alpha \tag{4}$$

But since f is integrable on $[a, c]$ and $[c, b]$, we have

$$\int_a^b f d\alpha \geq \int_a^c f d\alpha + \int_c^b f d\alpha \tag{5}$$

The relation (3) similarly yields

$$\int_a^b fd\alpha \le \int_a^c fd\alpha + \int_c^b fd\alpha \qquad (6)$$

Hence (5) and (6) imply that

$$\int_a^b fd\alpha = \int_a^c fd\alpha + \int_c^b fd\alpha .$$

Theorem 6. If $f \in \Re(\alpha)$, then

(i) $cf \in \Re(\alpha)$ and $\int_a^b (cf)d\alpha = c\int_a^b fd\alpha$, for every constant c,

(ii) If in addition $|f(x)| \le K$ on $[a,b]$, then

$$\left|\int_a^b fd\alpha\right| \le K[\alpha(b)-\alpha(a)] .$$

Proof.(i) Let $f \in \Re(\alpha)$ and let $g = cf$. Then

$$U(P,g,\alpha) = \sum_{i=1}^{n} M_i^{'}\Delta\alpha_i = \sum_{i=1}^{n} cM_i\Delta\alpha_i$$

$$= c\sum_{i=1}^{n} M_i\Delta\alpha_i$$

$$= cU(P,f,\alpha)$$

Similarly

$$L(P,g,\alpha) = cL(P,f,\alpha)$$

Since $f \in \Re(\alpha)$, $\exists$ a partition P such that for every $\in > 0$,

$$U(P,f,\alpha) - L(P,f,\alpha) < \frac{\in}{c}$$

Hence

$$U(P,g,\alpha) - L(P,g,\alpha) = c[U(P,f,\alpha) - L(P,f,\alpha)]$$

$$< c.\frac{\in}{c} = \in .$$

Hence $g = cf \in \Re(\alpha)$.

Further, since $U(P,f,\alpha) < \int_a^b fd\alpha + \frac{\in}{2}$,

$$\int_a^b g d\alpha \le U(P,g,\alpha) = cU(P,f,\alpha)$$

$$< c\left(\int_a^b f d\alpha + \frac{\in}{2}\right)$$

Since $\in$ is arbitrary

$$\int_a^b g d\alpha \le c\int_a^b f d\alpha$$

Replacing f by $-f$, we get

$$\int_a^b g d\alpha \ge c\int_a^b f d\alpha$$

Hence $\int_a^b (cf) d\alpha = c\int_a^b f d\alpha$.

(ii) If M and m are bounds of $f \in \Re(\alpha)$ on $[a,b]$, then it follows that

$$m[\alpha(b)-\alpha(a)] \le \int_a^b f d\alpha \le M[\alpha(b)-\alpha(a)] \text{ for } b \ge a \qquad (1).$$

In fact, if $a = b$, then (1) is trivial. If $b > a$, then for any partition P, we have

$$m[\alpha(b)-\alpha(a)] \le \sum_{i=1}^{n} m_i \Delta\alpha_i = L(P,f,\alpha)$$

$$\le \int_a^b f d\alpha$$

$$\le U(P,f,\alpha) = \sum M_i \Delta\alpha_i$$

$$\le M[\alpha(b)-\alpha(a)]$$

which yields

$$m[\alpha(b)-\alpha(a)] \le \int_a^b f d\alpha \le M[\alpha(b)-\alpha(a)] \qquad (2)$$

Since $|f(x)| \le K$ for all $x \in (a,b)$, we have

$$-K \le f(x) \le K$$

so if m and M are the bounds of f in (a,b),

$$-K \le m \le f(x) \le M \le K \text{ for all } x \in (a,b).$$

If $b \ge a$, then $\alpha(b)-\alpha(a) \ge 0$ and we have by (2)

$$-K[\alpha(b)-\alpha(a)] \le m[\alpha(b)-\alpha(a)] \le \int_a^b f d\alpha$$

$$\le M[\alpha(b)-\alpha(a)] \le K[\alpha(b)-\alpha(a)]$$

Hence

$$\left|\int_a^b f d\alpha\right| \le K[\alpha(b)-\alpha(a)].$$

Theorem 7. Suppose $f \in \Re(\alpha)$ on $[a,b]$, $m \le f \le M$, ϕ is continuous on $[m,M]$ and $h(x)=\phi[f(x)]$ on $[a,b]$. Then $h \in \Re(\alpha)$ on $[a,b]$.

Proof. Let $\in > 0$. Since ϕ is continuous on closed and bounded interval $[m,M]$, it is uniformly continuous on $[m,M]$. Therefore there exists $\delta > 0$ such that $\delta < \in$ and

$$|\phi(s)-\phi(t)| < \in \text{ if } |s-t| \le \delta,\ s,t \in [m,M].$$

Since $f \in \Re(\alpha)$, there is a partition $P=\{x_0, x_1, ..., x_n\}$ of $[a,b]$ such that

$$U(P,f,\alpha)-L(P,f,\alpha) < \delta^2 \qquad (1).$$

Let M_i, m_i and M_i^*, m_i^* be the lub, glb of $f(x)$ and $\phi(x)$ respectively in $[x_{i-1}, x_i]$. Divide the number 1,2,……,n into two classes:

$$i \in A \text{ if } M_i - m_i < \delta$$

and

$$i \in B \text{ if } M_i - m_i \ge \delta.$$

For $i \in A$, our choice of δ implies that $M_i^* - m_i^* \le \in$. Also, for $i \in B$, $M_i^* - m_i^* \le 2k$ where $k = \text{lub}|\phi(t)|$, $t \in [m,M]$. Hence, using (1), we have

$$\delta \sum_{i\in B} \Delta\alpha_i \le \sum_{i \in B} (M_i - m_i)\Delta\alpha_i < \delta^2 \qquad (2)$$

so that $\sum_{i\in B} \Delta\alpha_i < \delta$. Then we have

$$U(P,h,\alpha)-L(P,h,\alpha) = \sum_{i\in A}(M_i^* - m_i^*)\Delta\alpha_i + \sum_{i\in B}(M_i^* - m_i^*)\Delta\alpha_i$$

$$\le \in [\alpha(b)-\alpha(a)] + 2k\delta$$

$$\le [[\alpha(b)-\alpha(a)] + 2k] \in$$

Since $\in$ was arbitrary,

$$U(P,h,\alpha)-L(P,h,\alpha)<\in^*,\ \in^*>0\ .$$

Hence $h\in f(\alpha)$.

Theorem 8. If $f\in\Re(\alpha)$ and $g\in\Re(\alpha)$ on $[a,b]$, then $fg\in\Re(\alpha)$, $|f|\in\Re(\alpha)$ and

$$\left|\int_a^b fd\alpha\right|\le\int_a^b|f|\,d\alpha\ .$$

Proof. Let ϕ be defined by $\phi(t)=t^2$ on $[a,b]$. Then $h(x)=\phi[f(x)]=f^2\in\Re(\alpha)$ by Theorem7(in section 1.3.2). Also

$$fg=\frac{1}{4}[(f+g)^2-(f-g)^2]\ .$$

Since $f,g\in\Re(\alpha)$, $f+g\in\Re(\alpha)$, $f-g\in\Re(\alpha)$. Then $(f+g)^2$ and $(f-g)^2\in\Re(\alpha)$ and so their difference multiplied by $\frac{1}{4}$ also belong to $\Re(\alpha)$ proving that $fg\in\Re(\alpha)$.

If we take $\phi(f)=|t|$, again Theorem 7 implies that $|f|\in\Re(\alpha)$. We choose $c=\pm1$ so that

$$c\int fd\alpha\ge 0$$

Then

$$\left|\int fd\alpha\right|=c\int fd\alpha=\int cfd\alpha\le\int|f|\,d\alpha$$

because $cf\le|f|$.

1.3.3. Riemann-Stieltjes integral as limit of sums. In this section, we shall show that Riemann-Stieltjes integral $\int fd\alpha$ can be considered as the limit of a sequence of sums in which M_i,m_i involved in the definition of $\int fd\alpha$ are replaced by the values of f.

Definition 1. Let $P=\{a=x_0,x_1,........,x_n=b\}$ be a partition of $[a,b]$ and let points $t_1,t_2,......,t_n$ be such that $t_i\in[x_{i-1},x_i]$. Then the sum

$$S(P,f,\alpha)=\sum_{i=1}^{n}f(t_i)\Delta\alpha_i$$

is called a **Riemann-Stieltjes sum of f with respect to α .**

Definition 2. We say that

$$\lim_{|P|\to 0} S(P,f,\alpha)=A$$

If for every $\in>0$, there exists a $\delta>0$ such that $|P|<\delta$ implies

$$|S(P,f,\alpha)-A|<\in.$$

Theorem 1. If $\lim_{|P|\to 0} S(P,f,\alpha)$ exists, then $f\in\Re(\alpha)$ and

$$\lim_{|P|\to 0} S(P,f,\alpha)=\int_a^b f d\alpha .$$

Proof. Suppose $\lim_{|P|\to 0} S(P,f,\alpha)$ exists and is equal to A. Then given $\in>0$ there exists a $\delta>0$ such that $|P|<\delta$ implies

$$|S(P,f,\alpha)-A|<\frac{\in}{2}$$

or

$$A-\frac{\in}{2}<S(P,f,\alpha)<A+\frac{\in}{2} \qquad (1)$$

If we choose partition P satisfying $|P|<\delta$ and if we allow the points t_i to range over $[x_{i-1},x_i]$, taking lub and glb of the numbers $S(P,f,\alpha)$ obtained in this way, the relation (1) gives

$$A-\frac{\in}{2}\le L(P,f,\alpha)\le U(P,f,\alpha)\le A+\frac{\in}{2}$$

and so

$$U(P,f,\alpha)-L(P,f,\alpha)<\frac{\in}{2}+\frac{\in}{2}=\in$$

Hence $f\in\Re(\alpha)$. Further

$$A-\frac{\in}{2}\le L(P,f,\alpha)\le\int f d\alpha\le U(P,f,\alpha)\le A+\frac{\in}{2}$$

which yields

$$A-\frac{\in}{2}\le\int f d\alpha\le A+\frac{\in}{2}$$

or

$$\int f d\alpha=A=\lim_{|P|\to 0} S(P,f,\alpha).$$

Theorem 2. If

(i) f is continuous, then

$$\lim_{|P|\to 0} S(P,f,\alpha)=\int_a^b f d\alpha$$

(ii) $f\in\Re(\alpha)$ and α is continuous on $[a,b]$, then

$$\lim_{|P|\to 0} S(P,f,\alpha)=\int_a^b f d\alpha .$$

Proof. Part (i) is already proved in Theorem 1(ii) of section 1.3.2 of this unit.

(ii) Let $f\in\Re(\alpha)$, α be continuous and $\in>0$. Then there exists a partition P^* such that

$$U(P^*,f,\alpha)<\int f d\alpha+\frac{\in}{4} \qquad (1)$$

Now, α being uniformly continuous, there exists $\delta_1>0$ such that for any partition P of $[a,b]$ with $|P|<\delta_1$, we have

$$\Delta\alpha_i=\alpha(x_i)-\alpha(x_{i-1})<\frac{\in}{4Mn} \text{ for all i}$$

where n is the number of intervals into which P^* divides $[a,b]$. Consider the sum $U(P,f,\alpha)$. Those intervals of P which contains a point of P^* in their interior contribute no more than:

$$(n-1)\max\Delta\alpha_i.M<\frac{(n-1)\in M}{4Mn}<\frac{\in}{4} \text{ to } U(P^*,f,\alpha) \qquad (2)$$

Then (1) and (2) yield

$$U(P,f,\alpha)<\int f d\alpha+\frac{\in}{2} \qquad (3)$$

for all P with $|P|<\delta_1$.

Similarly, we can show that there exists a $\delta_2>0$ such that

$$L(P,f,\alpha)>\int f d\alpha-\frac{\in}{2} \qquad (4)$$

for all P with $|P|<\delta_2$.

Taking $\delta=\min\{\delta_1,\delta_2\}$, it follows that (2) and (3) hold for every P such that $|P|<\delta$.

Since

$$L(P,f,\alpha)\le S(P,f,\alpha)\le U(P,f,\alpha)$$

(3) and (4) yield

$$S(P, f, \alpha) < \int f d\alpha + \frac{\in}{2}$$

and

$$S(P, f, \alpha) > \int f d\alpha - \frac{\in}{2}$$

Hence

$$\left| S(P, f, \alpha) - \int f d\alpha \right| < \frac{\in}{2}$$

for all P such that $|P| < \delta$ and so

$$\lim_{|P| \to 0} S(P, f, \alpha) = \int f d\alpha$$

This completes the proof of the theorem.

The **Abel's Transformation (Partial Summation Formula)** for sequences reads as follows:

Let $< a_n >$ and $< b_n >$ be sequences and let

$$A_n = a_0 + a_1 + ... + a_n \qquad (A_{-1} = 0),$$

then

$$\sum_{n=p}^{q} a_n b_n = \sum_{n=p}^{q-1} A_n (b_n - b_{n+1}) + A_q b_q - A_{p-1} b_p$$

1.4 Integration and Differentiation. In this section, we show that integration and differentiation are inverse operations.

Definition 1. If $f \in \Re$ on $[a, b]$, then the function F defined by

$$F(t) = \int_a^t f(x) dx, \ t \in [a, b]$$

is called the "**Integral Function**" of the function f.

Theorem 1. If $f \in \Re$ on $[a, b]$, then the integral function F of f is continuous on $[a, b]$.

Proof. We have

$$F(t) = \int_a^t f(x) dx$$

Since $f \in \Re$, it is bounded and therefore there exists a number M such that for all x in $[a, b]$, $|f(x)| \le M$.

Let $\in$ be any positive number and c be any point of $[a,b]$. Then

$$F(c)=\int_a^c f(x)dx\,,\ F(c+h)=\int_a^{c+h} f(x)dx$$

Therefore

$$\begin{aligned}|F(c+h)-F(c)| &= \left|\int_a^{c+h} f(x)dx-\int_a^c f(x)dx\right| \\ &= \left|\int_c^{c+h} f(x)dx\right| \\ &\le M\,|h| \\ &< \in \quad \text{if } |h|<\frac{\in}{M}.\end{aligned}$$

Thus $|(c+h)-c|<\delta=\dfrac{\in}{M}$ implies $|F(c+h)-F(c)|<\in$. Hence F is continuous at any point $c\in[a,b]$ and is so continuous in the interval $[a,b]$.

Theorem 2. If f is continuous on $[a,b]$, then the integral function F is differentiable and

$$F'(x_0)=f(x_0)\,,\ x\in[a,b].$$

Proof. Let f be continuous at x_0 in $[a,b]$. Then there exists $\delta>0$ for every $\in>0$ such that

$$|f(t)-f(x_0)|<\in \qquad (1)$$

whenever $|t-x_0|<\delta$. Let $x_0-\delta<s\le x_0\le t<x_0+\delta$ and $a\le s<t\le b$, then

$$\begin{aligned}\left|\frac{F(t)-F(s)}{t-s}-f(x_0)\right| &= \left|\frac{1}{t-s}\int_s^t f(x)dx-f(x_0)\right| \\ &= \left|\frac{1}{t-s}\int_s^t f(x)dx-\frac{1}{t-s}\int_s^t f(x_0)dx\right| \\ &= \left|\frac{1}{t-s}\int_s^t [f(x)-f(x_0)]dx\right| \le \frac{1}{t-s}\left|\int_s^t [f(x)-f(x_0)]dx\right|<\in,\end{aligned}$$

(using (1)).

Hence $F'(x_0)=f(x_0)$. This completes the proof of the theorem.

Definition 2. A derivable function F such that F' is equal to a given function f in $[a,b]$ is called **Primitive of f .**

Thus the above theorem asserts that "Every continuous function f possesses a Primitive, viz the integral function $\int_a^t f(x)dx$."

Furthermore, the continuity of a function is not necessary for the existence of primitive. In other words, the function possessing primitive is not necessary continuous. For example, consider the function f on $[0,1]$ defined by

$$f(x)=\begin{cases} 2x\sin\frac{1}{x}-\cos\frac{1}{x}, x\neq 0 \\ 0, x=0 \end{cases}$$

It has primitive

$$F(x)=\begin{cases} x^2\sin\frac{1}{x}, x\neq 0 \\ 0, x=0 \end{cases}$$

Clearly $F^{'}(x)=f(x)$ but $f(x)$ is not continuous at $x=0$, i.e., f is not continuous in $[0,1]$.

1.5 Fundamental Theorem of the Integral Calculus

Theorem 1 (Fundamental Theorem of the Integral Calculus). If f ∈ R on $[a,b]$ and if there is a differential function F on $[a,b]$ such that $F'=f$, then

$$\int_a^b f(x)dx = F(b)-F(a).$$

Proof. Let P be a partition of $[a,b]$ and choose $t_i, (i=1,2,......,n)$ such that $x_{i-1}\leq t_i \leq x_i$. Then, by Lagrange's Mean Value Theorem, we have

$$F(x_i)-F(x_{i-1})=(x_i-x_{i-1})F'(t_i)=(x_i-x_{i-1})f(t_i) \qquad \text{(Since } F'=f \text{).}$$

Further

$$\begin{aligned} F(b)-F(a) &= \sum_{i=1}^{n}[F(x_i)-F(x_{i-1})] \\ &= \sum_{i=1}^{n} f(t_i)(x_i-x_{i-1}) \\ &= \sum_{i=1}^{n} f(t_i)\Delta x_i \end{aligned}$$

and the last sum tends to $\int_a^b f(x)dx$ as $|P|\to 0$, by theorem 1 of section 1.3.3, taking $\alpha(x)=x$. Hence

$$\int_a^b f(x)dx = F(b) - F(a).$$

This completes the proof of the theorem.

The next theorem tells us that the symbols $d\alpha(x)$ can be replaced by $\alpha'(x)dx$ in the Riemann-Stieltjes integral $\int_a^b f(x)d\alpha(x)$. This is the situation in which Riemann-Stieltjes integral reduces to Riemann integral.

Theorem 2. If $f \in \Re$ and $\alpha' \in \Re$ on $[a,b]$, then $f \in \Re(\alpha)$ and

$$\int_a^b f d\alpha = \int_a^b f(x)\alpha'(x)dx$$

Proof. Since $f \in \Re$, $\alpha' \in \Re$, it follows that their product $f\alpha' \in \Re$. Let $\in > 0$ be given. Choose M such that $|f| \le M$. Since $f\alpha' \in \Re$ and $\alpha' \in \Re$, using Theorem 2(ii) of section 1.3.3 for integrator as x, we have

$$\left|\sum f(t_i)\alpha'(t_i)\Delta x_i - \int f\alpha'\right| < \in \qquad (1).$$

if $|P| < \delta_1$ and $x_{i-1} \le t_i \le x_i$ and

$$\left|\sum \alpha'(t_i)\Delta x_i - \int \alpha'\right| < \in \qquad (2).$$

if $|P| < \delta_2$ and $x_{i-1} \le t_i \le x_i$. Letting t_i vary in (2), we have

$$\left|\sum \alpha'(s_i)\Delta x_i - \int \alpha'\right| < \in \qquad (3).$$

if $|P| < \delta_2$ and $x_{i-1} \le s_i \le x_i$. From (2) and (3) it follows that

$$\left|\sum \alpha'(t_i)\Delta x_i - \int \alpha' + \int \alpha' - \sum \alpha'(s_i)\Delta x_i\right|$$

$$\le \left|\sum \alpha'(t_i)\Delta x_i - \int \alpha'\right| + \left|\sum \alpha'(s_i)\Delta x_i - \int \alpha'\right|$$

$$< \in + \in = 2\in$$

or

$$\sum \left|\alpha'(t_i) - \alpha'(s_i)\right| \Delta x_i < 2\in \qquad (4).$$

if $|P| < \delta_2$ and $x_{i-1} \le t_i \le x_i$, $x_{i-1} \le s_i \le x_i$.

Now choose a partition P so that $|P| < \delta = \min\{\delta_1, \delta_2\}$ and choose $t_i \in [x_{i-1}, x_i]$. By Mean Value Theorem,

$$\Delta\alpha_i = \alpha(x_i) - \alpha(x_{i-1}) = \alpha'(s_i)(x_i - x_{i-1})$$
$$= \alpha'(s_i)\Delta x_i$$

Then, we have

$$\sum f(t_i)\Delta\alpha_i = \sum f(t_i)\alpha'(t_i)\Delta x_i + \sum f(t_i)[\alpha'(s_i) - \alpha'(t_i)]\Delta x_i \qquad (5).$$

Thus, by (1) and (4), it follows that

$$\left|\sum f(t_i)\Delta\alpha_i - \int f\alpha'\right| = \left|\sum f(t_i)\alpha'(t_i)\Delta x_i - \int f\alpha' + \sum f(t_i)[\alpha'(s_i) - \alpha'(t_i)]\Delta x_i\right|$$
$$< \in + 2 \in M = \in (1 + 2M)$$

Hence

$$\lim_{|P|\to 0} \sum f(t_i)\Delta x_i = \int_a^b f(x)\alpha'(x)dx$$

or

$$\int_a^b f d\alpha = \int_a^b f(x)\alpha'(x)dx .$$

Example 1. Evaluate (i) $\int_0^2 x^2 dx^2$, (ii) $\int_0^2 [x] dx^2$.

Solution. We know that

$$\int_a^b f d\alpha = \int_a^b f(x)\alpha'(x)dx$$

Therefore

$$\int_0^2 x^2 dx^2 = \int_0^2 x^2(2x)dx = \int_0^2 2x^3 dx$$
$$= 2\left|\frac{x^4}{4}\right|_0^2 = 8 .$$

and

$$\int_0^2 [x] dx^2 = \int_0^2 [x] 2x dx$$
$$= \int_0^1 [x] 2x dx + \int_1^2 [x] 2x dx$$
$$= 0 + \int_1^2 2x dx = 0 + 2\left|\frac{x^2}{2}\right|_1^2$$

=0+3 = 3.

We now establish a connection between the integrand and the integrator in a Riemann-Stieltjes integral. We shall show that existence of $\int f d\alpha$ implies the existence of $\int \alpha df$.

We recall that Abel's transformation (Partial Summation Formula) for sequences reads as follows:

"Let $<a_n>$ and $<b_n>$ be two sequences and let $A_n = a_0 + a_1 + ... + a_n$ $(A_{-1} = 0)$. Then

$$\sum_{n=p}^{q} a_n b_n = \sum_{n=p}^{q-1} A_n (b_n - b_{n+1}) + A_q b_q - A_{p-1} b_p \qquad (*)."$$

1.5.1 Theorem (Integration by parts). If $f \in \Re(\alpha)$ on $[a,b]$, then $\alpha \in \Re(f)$ on $[a,b]$ and

$$\int f(x) d\alpha(x) = f(b)\alpha(b) - f(a)\alpha(a) - \int \alpha(x) df(x)$$

(Due to analogy with (*), the above expression is also known as **Partial Integration Formula**).

Proof. Let $P = \{a = x_0, x_1, ..., x_n = b\}$ be a partition of $[a,b]$. Choose $t_1, t_2, ..., t_n$ such that $x_{i-1} \le t_i \le x_i$ and take $t_0 = a$, $t_{n+1} = b$. Suppose Q is the partition $\{t_1, t_2, ..., t_{n+1}\}$ of $[a,b]$. By

partial summation, we have

$$S(P, f, \alpha) = \sum_{i=1}^{n} f(t_i)[\alpha(x_i) - \alpha(x_{i-1})] = f(b)\alpha(b) - f(a)\alpha(a) - \sum_{i=1}^{n+1} \alpha(x_{i-1})[f(t_i) - f(t_{i-1})]$$

$$= f(b)\alpha(b) - f(a)\alpha(a) - S(Q, \alpha, f)$$

since $t_{i-1} \le x_{i-1} \le t_i$. If $|P| \to 0$, $|Q| \to 0$, then

$$S(P, f, \alpha) \to \int f d\alpha \text{ and } S(Q, \alpha, f) \to \int \alpha df .$$

Hence

$$\int f d\alpha = f(b)\alpha(b) - f(a)\alpha(a) - \int \alpha df .$$

1.5.2 Mean Value Theorems for Riemann-Stieltjes Integrals. In this section, we establish Mean Value Theorems which are used to get estimate value of an integral rather than its exact value.

Theorem 1.5.2(a). (First Mean Value Theorem for Riemann-Stieltjes Integral). If f is continuous and real valued and α be is monotonically increasing on $[a,b]$, then there exists a point x in $[a,b]$ such that

$$\int_a^b f d\alpha = f(x)[\alpha(b) - \alpha(a)].$$

Proof. If $\alpha(a) = \alpha(b)$, the theorem holds trivially, both sides being 0 in that case (α become constant and so $d\alpha = 0$). Hence we assume that $\alpha(a) < \alpha(b)$. Let

$$M = \text{lub}\, f(x),\ m = \text{glb}\, f(x).\ \ a \le x \le b$$

Then

$$m \le f(x) \le M$$

or

$$m[\alpha(b)-\alpha(a)] \le \int f d\alpha \le M[\alpha(b)-\alpha(a)]$$

Hence there exists some c satisfying $m \le c \le M$ such that

$$\int_a^b f d\alpha = c[\alpha(b)-\alpha(a)]$$

Since f is continuous, there is a point $x \in [a,b]$ such that $f(x)=c$ and so we have

$$\int_a^b f(x)d\alpha(x) = f(x)[\alpha(b)-\alpha(a)]$$

This completes the proof of the theorem.

Theorem 1.5.2(b) (Second Mean Value Theorem for Riemann-Stieltjes Integral). Let f be monotonic and α be real and continuous. Then there is a point $x \in [a,b]$ such that

$$\int_a^b f d\alpha = f(a)[\alpha(x)-\alpha(a)] + f(b)[\alpha(b)-\alpha(x)]$$

Proof. By Partial Integration Formula, we have

$$\int_a^b f d\alpha = f(b)\alpha(b) - f(a)\alpha(a) - \int_a^b \alpha df$$

The use of First Mean Value Theorem for Riemann-Stieltjes Integral yields that there is x in $[a,b]$ such that

$$\int_a^b \alpha df = \alpha(x)[f(b)-f(a)]$$

Hence, for some $x \in [a,b]$, we have

$$\int_a^b f d\alpha = f(b)\alpha(b) - f(a)\alpha(a) - \alpha(x)[f(b)-f(a)]$$

$$= f(a)[\alpha(x)-\alpha(a)] + f(b)[\alpha(b)-\alpha(x)]$$

which proves the theorem.

1.5.3 We discuss now change of variable. In this direction we prove the following result.

Theorem 1. Let f and ϕ be continuous on $[a,b]$. If ϕ is strictly increasing on $[\alpha,\beta]$, where $a=\phi(\alpha)$, $b=\phi(\beta)$, then

$$\int_a^b f(x)dx = \int_\alpha^\beta f(\phi(y))d\phi(y)$$

(this corresponds to change of variable in $\int_a^b f(x)dx$ by taking $x = \phi(y)$).

Proof. Since ϕ is strictly monotonically increasing, it is invertible and so

$$\alpha = \phi^{-1}(a)\,,\ \ \beta = \phi^{-1}(b)\,.$$

Let $P = \{a = x_0, x_1, ..., x_n = b\}$ be any partition of $[a,b]$ and $Q = \{\alpha = y_0, y_1, ..., y_n = \beta\}$ be the corresponding partition of $[\alpha, \beta]$, where $y_i = \phi^{-1}(x_i)$. Then

$$\begin{aligned} \Delta x_i &= x_i - x_{i-1} \\ &= \phi(y_i) - \phi(y_{i-1}) \\ &= \Delta\phi_i\,. \end{aligned}$$

Let for any $c_i \in \Delta x_i$, $d_i \in \Delta y_i$, where $c_i \in \phi(d_i)$. Putting $g(y) = f[\phi(y)]$, we have

$$\begin{aligned} S(P,f) &= \sum_{i=1}^n f(c_i)\Delta x_i \qquad (1) \\ &= \sum_i f(\phi(d_i))\Delta\phi_i \\ &= \sum_i g(d_i)\Delta\phi_i \\ &= S(Q,g,\phi) \end{aligned}$$

Continuity of f implies that $S(P,f) \to \int_a^b f(x)dx$ as $|P| \to 0$ and continuity of g implies that

$$S(Q,g,\phi) \to \int_\alpha^\beta g(y)d\phi \text{ as} |Q| \to 0\,.$$

Since uniform continuity of ϕ on $[a,b]$ implies that $|Q| \to 0$ as $|P| \to 0$. Hence letting $|P| \to 0$ in (1), we have

$$\int_a^b f(x)dx = \int_\alpha^\beta g(y)d\phi = \int_\alpha^\beta f(\phi(y))d\phi(y)$$

This completes the proof of the theorem.

1.6 Integration of Vector –Valued Functions. Let $f_1, f_2, ..., f_k$ be real valued functions defined on $[a,b]$ and let $f = (f_1, f_2, ..., f_k)$ be the corresponding mapping of $[a,b]$ into R^k.

Let α be monotonically increasing function on $[a,b]$. If $f_i \in \Re(\alpha)$ for $i = 1, 2, ..., k$, we say that $f \in \Re(\alpha)$ and then the integral of f is defined as

$$\int_a^b f d\alpha = (\int_a^b f_1 d\alpha, \int_a^b f_2 d\alpha, ..., \int_a^b f_k d\alpha).$$

Thus $\int_a^b f d\alpha$ is the point in R^k whose ith coordinate is $\int_a^b f_i d\alpha$.

It can be shown that if $f \in \Re(\alpha)$, $g \in \Re(\alpha)$,

then

(i) $\int_a^b (f+g) d\alpha = \int_a^b f d\alpha + \int_a^b g d\alpha$

(ii) $\int_a^b f d\alpha = \int_a^c f d\alpha + \int_c^b f d\alpha$, $a < c < b$.

(iii) if $f \in \Re(\alpha_1)$, $f \in \Re(\alpha_2)$, then $f \in \Re(\alpha_1 + \alpha_2)$

and

$$\int_a^b f d(\alpha_1 + \alpha_2) = \int_a^b f d\alpha_1 + \int_a^b f d\alpha_2$$

To prove these results, we have to apply earlier results to each coordinate of f. Also, fundamental theorem of integral calculus holds for vector valued function f. We have

Theorem 1. If f and F map $[a,b]$ into $\Re^k$, if $f \in \Re(\alpha)$ if $F' = f$, then

$$\int_a^b f(t) dt = F(b) - F(a)$$

Theorem 2. If f maps $[a,b]$ into R^k and if $f \in R(\alpha)$ for some monotonically increasing function α on $[a,b]$, then $|f| \in R(\alpha)$ and

$$\left| \int_a^b f d\alpha \right| \le \int_a^b |f| d\alpha.$$

Proof. Let

$$f = (f_1, ..., f_k).$$

Then

$$|f| = (f_1^2 + ... + f_k^2)^{1/2}$$

Since each $f_i \in R(\alpha)$, the function $f_i^2 \in R(\alpha)$ and so their sum $f_1^2 + ... + f_k^2 \in R(\alpha)$. Since x^2 is a continuous function of x, the square root function is continuous on $[0, M]$ for every real M. Therefore $|f| \in R(\alpha)$.

Now, let $y = (y_1, y_2, ...y_k)$, where $y_i = \int f_i d\alpha$, then

$$y = \int f d\alpha$$

and

$$|y|^2 = \sum_i y_i^2 = \sum y_i \int f_i d\alpha$$

$$= \int (\sum y_i f_i) d\alpha$$

But, by Schwarz inequality

$$\left|\sum y_i f_i(t)\right| \le |y||f(t)|, \ (a \le t \le b)$$

Then

$$|y|^2 \le |y| \int |f| d\alpha \qquad (1)$$

If $y = 0$, then the result follows. If $|y| \ne 0$, then divide (1) by $|y|$ and get

$$|y| \le \int |f| d\alpha$$

or
$$\left|\int_a^b f d\alpha\right| \le \int_a^b |f| d\alpha .$$

1.7 Rectifiable Curves. The aim of this section is to consider application of results studied in this chapter to geometry.

Definition 1. A continuous mapping γ of an interval $[a,b]$ into R^k is called a curve in R^k.

If $\gamma : [a,b] \to R^k$ is continuous and one-to-one, then it is called an arc.

If for a curve $\gamma : [a,b] \to R^k$,

$$\gamma(a) = \gamma(b)$$

but

$$\gamma(t_1) \ne \gamma(t_2)$$

for every other pair of distinct points t_1, t_2 in $[a,b]$, then the curve γ is called a simple closed curve.

Definition 2. Let $f : [a,b] \to R^k$ be a map. If $P = \{x_0, x_1, ..., x_n\}$ is a partition of $[a,b]$, then

$$V(f, a, b) = \text{lub} \sum_{i=1}^{n} |f(x_i) - f(x_{i-1})|,$$

where the lub is taken over all possible partitions of $[a,b]$, is called total variation of f on $[a,b]$. The function f is said to be of bounded variation on $[a,b]$ if $V(f,a,b) < +\infty$.

Definition 3. A curve $\gamma:[a,b]\to R^k$ is called rectifiable if γ is of bounded variation. The length of a rectifiable curve γ is defined as total variation of γ, i.e., $V(\gamma,a,b)$. Thus length of rectifiable curve

$\gamma = \text{lub}\sum_{i=1}^{n}|\gamma(x_i)-\gamma(x_{i-1})|$ for the partition $(a = x_0 < x_1 < .... < x_n = b)$.

The i^th^ term $|\gamma(x_i)-\gamma(x_{i-1})|$ in this sum is the distance in R^k between the points $\gamma(x_{i-1})$ and $\gamma(x_i)$. Further $\sum_{i=1}^{n}|\gamma(x_i)-\gamma(x_{i-1})|$ is the length of a polygon whose vertices are at the points $\gamma(x_0),\gamma(x_1),....,\gamma(x_n)$. As the norm of our partition tends to zero, then those polygons approach the range of γ more and more closely.

Theorem 1. Let γ be a curve in R^k. If γ' is continuous on $[a,b]$, then γ is rectifiable and has length

$$\int_a^b |\gamma'(t)|\,dt .$$

Proof. It is sufficient to show that $\int|\gamma'| = V(\gamma,a,b)$. So, let $\{x_0,....,x_n\}$ be a partition of $[a,b]$.

Using Fundamental Theorem of Calculus for vector valued function, we have

$$\begin{aligned}\sum_{i=1}^{n}|\gamma(x_i)-\gamma(x_{i-1})| &= \sum_{i=1}^{n}\left|\int_{x_{i-1}}^{x_i}\gamma'(t)dt\right|\\ &\le \sum_{i=1}^{n}\int_{x_{i-1}}^{x_i}|\gamma'(t)|dt\\ &= \int_a^b|\gamma'(t)|dt\end{aligned}$$

Thus

$$V(\gamma,a,b)\le\int|\gamma'| \qquad (1).$$

To prove the reverse inequality, let $\in$ be a positive number. Since γ' is uniformly continuous on $[a,b]$, there exists $\delta>0$ such that

$$|\gamma'(s)-\gamma'(t)| < \in, \text{ if } |s-t|<\delta .$$

If mesh (norm) of the partition P is less than δ and $x_{i-1}\le t\le x_i$, then we have

$$|\gamma'(t)|\le|\gamma'(x_i)|+\in ,$$

so that

$$\int_{x_{i-1}}^{x_i} |\gamma'(t)|\,dt - \in \Delta x_i \leq |\gamma'(x_i)|\Delta x_i$$

$$= \left| \int_{x_{i-1}}^{x_i} [\gamma'(t) + \gamma'(x_i) - \gamma'(t)]dt \right|$$

$$\leq \left| \int_{x_{i-1}}^{x_i} \gamma'(t)dt \right| + \left| \int_{x_{i-1}}^{x_i} [\gamma'(x_i) - \gamma'(t)]dt \right|$$

$$\leq |\gamma(x_i) - \gamma(x_{i-1})| + \in \Delta x_i$$

Adding these inequalities for $i = 1, 2, ..., n$, we get

$$\int_a^b |\gamma'(t)|\,dt \leq \sum_{i=1}^{n} |\gamma(x_i) - \gamma(x_{i-1})| + \in (b-a)$$

$$= V(\gamma, a, b) + \in (b-a)$$

Since $\in$ is arbitrary, it follows that

$$\int_a^b |\gamma'(t)|\,dt \leq V(\gamma, a, b) \qquad (2).$$

Combining (1) and (2), we have

$$\int_a^b |\gamma'(t)|\,dt = V(\gamma, a, b)$$

Hence the length of γ is $\int_a^b |\gamma'(t)|\,dt$.

1.8 References

1. Walter Rudin, Principles of Mathematical Analysis (3rd edition) McGraw-Hill, Kogakusha, 1976, International Student Edition.
2. T. M. Apostol, Mathematical Analysis, Narosa Publishing House, New Delhi, 1974.
3. H.L. Royden, Real Analysis, Macmillan Pub. Co., Inc. 4th Edition, New York, 1993.
4. G. De Barra, Measure Theory and Integration, Wiley Eastern Limited, 1981.
5. R.R. Goldberg, Methods of Real Analysis, Oxford & IBH Pub. Co. Pvt. Ltd, 1976.
6. R. G. Bartle, The Elements of Real Analysis, Wiley International Edition, 2011.
7. S.C. Malik and Savita Arora, Mathematical Analysis, New Age International Limited, New Delhi, 2012.

UNIT – II

SEQUENCE AND SERIES OF FUNCTIONS

Structure

2.0 Introduction

In this unit, we will consider sequence and series of functions whose terms depend on a variable. Uniform convergence of sequence or series is a concept of great importance in its domain. With the help of tests for uniform convergence, we will naturally inquire how we can determine whether the given sequence or series does or does not converge uniformly in a given interval. The Weierstrass approximation theorem describes that every continuous function can be "uniformly approximated" by polynomials to within any degree of accuracy.

2.1 Unit Objectives

After going through this unit, one will be able to

- learn about pointwise and uniform convergence of sequence and series of functions
- examine uniform convergence through various tests for uniform convergence.
- study uniform convergence and continuity.
- understand importance of Weierstrass approximation theorem.

2.2 Sequence and Series of Functions

Let f_n be a real valued function defined on an interval I (or on a subset D of R) and for each $n \in N$, then <f_1, f_2,……… ,f_n,…….> is called a sequence of real valued functions on I. It is denoted by $\{f_n\}$ or <f_n>.

If $<f_n>$ is a sequence of real valued functions on an interval I, then $f_1+ f_2+\ldots\ldots\ldots +f_n+\ldots\ldots\ldots$.is called a series of real valued functions defined on I. This series is denoted by $\sum_{n=1}^{\infty} f_n$ or simply $\sum f_n$. That is, we shall consider sequences whose terms are functions rather than real numbers. These sequences are useful in obtaining approximations to a given function.

2.3 Pointwise and Uniform Convergence of Sequences of Functions

We shall study two different notations of convergence for a sequence of functions: Pointwise convergence and uniform convergence.

Definition 1. Let $A \subseteq R$ and suppose that for each $n \in N$ there is a function $f_n : A \to R$. Then $\langle f_n \rangle$ is called a **sequence of functions** on A. For each $x \in A$, this sequence gives rise to a sequence of real numbers, namely the sequence $\langle f_n(x) \rangle$.

Definition 2. Let $A \subseteq R$ and let $\langle f_n \rangle$ be a sequence of functions on A. Let $A_0 \subseteq A$ and suppose $f : A_0 \to R$. Then the sequence $\langle f_n \rangle$ is said to converge on A_0 to f if for each $x \in A_0$, the sequence $\langle f_n(x) \rangle$ converges to f(x) in R.

In such a case f is called the limit function on A_0 of the sequence $\langle f_n \rangle$.

When such a function f exists, we say that the sequence $\langle f_n \rangle$ is convergent on A_0 or that $\langle f_n \rangle$ **converges pointwise on** A_0 to f and we write $f(x) = \lim_{n\to\infty} f_n(x),\ x \in A_0$.

Similarly, if $\sum f_n(x)$ converges for every $x \in A_0$, and if $f(x) = \sum_{n=1}^{\infty} f_n(x),\ x \in A_0$. The function f is called the sum of the series $\sum f_n$.

The question arises: If each function of a sequence $\langle f_n \rangle$ has certain property, such as continuity, differentiability or integrability, then to what extent is this property transferred to the limit function? For example, if each function f_n is continuous at a point x_0, is the limit function f also continuous at x_0? In general, it is not true. Thus, pointwise convergence is not so strong concept which transfers above mentioned property to the limit function. Therefore some stronger methods of convergence are needed. One of these methods is the notion of uniform convergence:

We know that f_n is continuous at x_0 if $\lim_{x\to x_0} f_n(x) = f_n(x_0)$.On the other hand,

$$f \text{ is continuous at } x_0 \text{ if } \lim_{x\to x_0} f(x) = f(x_0) \qquad (1)$$

But (1) can be written as

$$\lim_{x\to x_0} \lim_{n\to\infty} f_n(x) = \lim_{n\to\infty} \lim_{x\to x_0} f_n(x) \qquad (2)$$

Thus our question of continuity reduces to "can we interchange the limit symbols in (2)?" or "Is the order in which limit processes are carried out immaterial ". The following examples show that the limit symbols cannot in general be interchanged.

Example 1. A sequence of continuous functions whose limit function is discontinuous:

Let

$$f_n(x) = \frac{x^{2n}}{1+x^{2n}}, x \in R,\ n = 1, 2, \ldots\ldots\ldots$$

We note that

$$\lim_{n\to\infty} f_n(x) = f(x) = \begin{cases} 0 & \text{if } \ |x|<1 \\ \frac{1}{2} & \text{if } \ |x|=1. \\ 1 & \text{if } \ |x|>1 \end{cases}$$

Each f_n is continuous on R but the limit function f is discontinuous at x = 1 and x = -1.

Example 2. A double sequence in which limit process cannot be interchanged:

For m =1, 2,….,

n = 1, 2,3,..., let us consider the double sequence

$$S_{mn} = \frac{m}{m+n}.$$

For every fixed n, we have

$$\lim_{m\to\infty} S_{mn} = 1$$

and so

$$\lim_{n\to\infty} \lim_{m\to\infty} S_{mn} = 1$$

On the other hand, for every fixed m, we have

$$\lim_{n\to\infty} S_{mn} = \lim_{n\to\infty} \frac{1}{1+\frac{n}{m}} = 0$$

and so

$$\lim_{m\to\infty} \lim_{n\to\infty} S_{mn} = 0$$

Hence $$\lim_{n\to\infty} \lim_{m\to\infty} S_{mn} \neq \lim_{m\to\infty} \lim_{n\to\infty} S_{mn}$$

Example 3. A sequence of functions for which limit of the integral is not equal to integral of the limit: Let

$$f_n(x) = n^2 x(1-x)^n, \quad x \in R, \quad n = 1, 2, \ldots\ldots$$

If $0 \le x \le 1$, then

$$f(x) = \lim_{n\to\infty} f_n(x) = 0$$

and so

$$\int_0^1 f(x)\, dx = 0.$$

But $$\int_0^1 f_n(x)\, dx = n^2 \int_0^1 x(1-x)^n\, dx$$

$$= \frac{n^2}{n+1} - \frac{n^2}{n+2}$$

$$= \frac{n^2}{(n+1)(n+2)}.$$

and so $$\lim_{n\to\infty} \int_0^1 f_n(x)\, dx = 1$$

Hence $$\lim_{n\to\infty} \int_0^1 f_n(x)\, dx \ne \int_0^1 (\lim_{n\to\infty} f_n(x))\, dx.$$

Example 4. A sequence of differentiable functions $\{f_n\}$ with limit 0 for which $\{f_n'\}$ diverges.

Let

$$f_n(x) = \frac{\sin nx}{\sqrt{n}} \quad \text{if} \quad x \in R, \; n = 1, 2,$$

Then $\lim_{n\to\infty} f_n(x) = 0$ for all x.

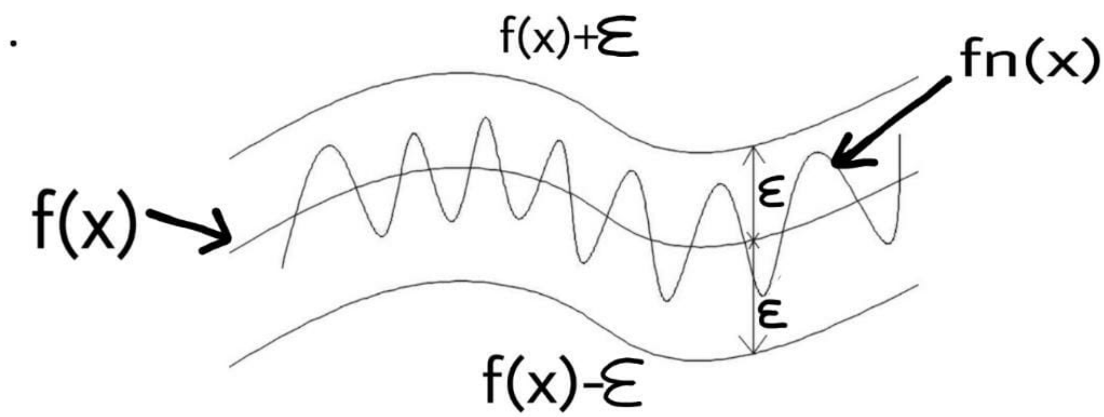
But $$f_n'(x) = \sqrt{n}\cos nx$$

and so $\lim_{n\to\infty} f_n'(x)$ does not exist for any x.

Definition 3. A sequence of functions $\{f_n\}$ is said to converge uniformly to a function f on a set E if for every ε> 0 there exists an integer N (depending only on ε) such that n > N implies

$$|f_n(x) - f(x)| < \varepsilon \qquad \text{for all } x \in E \qquad (*).$$

Geometrical Interpretation of uniform convergence:



If each term of the sequence $\langle f_n \rangle$ is real-valued, then the expression (*) can be written as

$f(x) - \varepsilon < f_n(x) < f(x) + \varepsilon$ for all n > N and for all $x \in E$.

This shows that the entire graph of f_n lies between a "band" of height 2ε situated symmetrically about the graph of f.

Definition 4. A series $\sum f_n(x)$ is said to converge uniformly on E if the sequence $\{S_n\}$ of partial sums defined by $S_n(x) = \sum_{i=1}^{n} f_i(x)$ converges uniformly on E.

Theorem 1. Every uniformly convergent sequence is pointwise convergent but not conversely.

Proof. Let $\{f_n\}$ be a sequence of functions which converges uniformly to f on E.

$\therefore$ For given $\varepsilon > 0$, there exists a positive integer N (depending only on ε) such that

$$|f_n(x) - f(x)| < \in \quad \text{for all } n > N \quad \ldots\ldots\ldots\ldots(1)$$

Since (1) is true for all $x \in E$.

$$|f_n(x) - f(x)| < \varepsilon \quad \text{for all } n > N$$

is true for every $x \in E$,

Hence f_n converges pointwise to f on E.

The converse is not true which is shown by following example.

Example 5. Consider the sequence $< f_n >$ defined by

$$f_n(x) = \frac{1}{nx+1}, 0 < x < 1$$

Then, $f(x) = \lim_{n\to\infty} f_n(x) = \lim_{n\to\infty} \frac{1}{nx+1} = 0$

Hence, $< f_n >$ converges pointwise to 0 for all $0 < x < 1$.

Let $\varepsilon > 0$ be given. Then for convergence, we have

$$|f_n(x) - f(x)| < \varepsilon \ , \ n > n_0$$

or $$\left|\frac{1}{nx+1} - 0\right| < \varepsilon \ , \ n > n_0.$$

or $$\frac{1}{nx+1} < \varepsilon.$$

or $$\frac{1}{nx} < \varepsilon$$

or $$nx > \frac{1}{\varepsilon}.$$

or $$n > \frac{1}{x\varepsilon}.$$

If n_0 is taken as integer greater than $\frac{1}{x\varepsilon}$, then

$$|f_n(x) - f(x)| < \varepsilon \quad \text{for all } n > n_0.$$

Since n_0 depends both on ε & x in (0,1), so f_n does not converge uniformly on (0,1).

Example 6. Consider the sequence $\langle S_n \rangle$ defined by $S_n(x) = \frac{1}{x+n}$ in any interval [a, b], a > 0. Then

$$S(x) = \lim_{n\to\infty} S_n(x) = \lim_{n\to\infty} \frac{1}{x+n} = 0$$

For the convergence, we must have

$$|S_n(x) - S(x)| < \varepsilon \quad , \quad n > n_0 \qquad (1)$$

or $$\left|\frac{1}{x+n} - 0\right| < \varepsilon \quad , \quad n > n_0$$

or $$\frac{1}{x+n} < \varepsilon$$

or $$x + n > \frac{1}{\varepsilon}$$

or $$n > \frac{1}{\varepsilon} - x$$

If we select n_0 as integer next higher to $\frac{1}{\varepsilon}$, then (1) is satisfied for m(integer) greater than $\frac{1}{\varepsilon}$ which does not depend on $x \in [a, b]$. Hence the sequence $\langle S_n \rangle$ is uniformly convergent to S(x) in [a, b].

Example 7. Consider the sequence $\langle f_n \rangle$ defined by

$$f_n(x) = \frac{x}{1+nx}, \quad x \geq 0$$

Then

$$f(x) = \lim_{n\to\infty} \frac{x}{1+nx} = 0 \text{ for all } x \geq 0.$$

Then $\langle f_n \rangle$ converges pointwise to 0 for all $x \geq 0$. Let $\varepsilon > 0$, then for convergence we must have

$$|f_n(x) - f(x)| < \varepsilon, n > n_0$$

or $$\left|\frac{x}{1+nx} - 0\right| < \varepsilon, n > n_0$$

$$\frac{x}{1+nx} < \varepsilon$$

$$x < \varepsilon + nx\varepsilon$$

$$nx\varepsilon > x - \varepsilon$$

$$n > \frac{x-\varepsilon}{x\varepsilon}$$

$$n > \frac{x}{x\varepsilon} = \frac{1}{\varepsilon}$$

If n_0 is taken as integer greater than $\frac{1}{\varepsilon}$, then

$$|f_n(x) - f(x)| < \varepsilon, \quad \text{for all } n > n_0 \text{ and for all } x \in [0, \infty)$$

Hence $\langle f_n \rangle$ converges uniformly to f on $[0, \infty)$.

Example 8. Consider the sequence $\langle f_n \rangle$ defined by

$$f_n(x) = x^n,\ 0 \le x \le 1$$

Then

$$f_n(x) = \lim_{n\to\infty} x^n = \begin{cases} 0 & \text{if } 0 \le x < 1 \\ 1 & \text{if } x = 1 \end{cases}$$

Let $\varepsilon > 0$ be given. Then for convergence, we must have

$$|f_n(x) - f(x)| < \varepsilon,\ n > n_0$$

or
$$x^n < \varepsilon$$

or
$$\left(\frac{1}{x}\right)^n > \frac{1}{\varepsilon}$$

or
$$n > \frac{\log\frac{1}{\varepsilon}}{\log\frac{1}{x}}.$$

Thus we should take n_0 to be an integer next higher to $\dfrac{\log\frac{1}{\varepsilon}}{\log\frac{1}{x}}$. If we take x = 1, then m does not exist.

Thus the sequence in question is not uniformly convergent to f in the interval which contains 1.

Definition 5 (Point of non–uniform convergence). A point which is such as the sequence is non – uniformly convergent in any interval containing that point is called a point of non–uniform convergence.

In the following example x = 0 is a point of non–uniform convergence.

Example 9. Consider the sequence $\langle f_n \rangle$ defined by $f_n(x) = \dfrac{nx}{1+n^2x^2},\ 0 \le x \le a$.

Then if x = 0, then $f_n(x) = 0$

and so $f(x) = \lim_{n\to\infty} f_n(x) = 0.$

If $x \ne 0$, then

$$f(x) = \lim_{n\to\infty} f_n(x) = \lim_{n\to\infty} \frac{nx}{1+n^2x^2} = 0.$$

Thus f is continuous at x = 0. For convergence, we must have

$$|f_n(x) - f(x)| < \varepsilon,\ n > n_0.$$

or $$\frac{nx}{1+n^2x^2} < \varepsilon.$$

or $$1+n^2x^2 - \frac{nx}{\varepsilon} > 0.$$

or $$nx > \frac{1}{2\varepsilon} + \frac{1}{2}\sqrt{\frac{1}{\varepsilon^2} - 4}.$$

Thus we can find an upper bound for n in any interval $0 < a \le x \le b$, but the upper bound is infinite if the interval includes 0. Hence the given sequence is non-uniformly convergent in any interval which includes the origin. So 0 is the point of non-uniform convergence for this sequence.

Example 10. Consider the sequence $\langle f_n \rangle$ defined by

$$f_n(x) = \tan^{-1} nx,\ \ 0 \le x \le a.$$

Then

$$f(x) = \lim_{n\to\infty} f_n(x) = \begin{cases} \dfrac{\pi}{2} & \text{if } x \ne 0 \\ 0 & \text{if } x = 0 \end{cases}.$$

Thus the function is discontinuous at x = 0.

For convergence, we must have for $\varepsilon > 0$,

$$\left|f_n(x) - f(x)\right| < \varepsilon,\ \ n > n_0$$

or $$\frac{\pi}{2} - \tan^{-1} nx < \varepsilon$$

or $$\cot^{-1} nx < \varepsilon$$

or $$nx > \frac{1}{\tan\varepsilon}$$

or $$n > \frac{1}{\tan\varepsilon}\left(\frac{1}{x}\right)$$

Thus no upper bound can be found for the function on the right if 0 is an end point of the interval. Hence the convergence is non-uniform in any interval which includes 0. So, here 0 is the point of non-uniform convergence.

Definition 6. A sequence $\{f_n\}$ is said to be **uniformly bounded** on E if there exists a constant $M > 0$ such that $\left|f_n(x)\right| \le M$ for all x in E and all n. The number M is called a uniform bound for $\{f_n\}$.

For example, the sequence $\langle f_n \rangle$ defined by $f_n(x) = \sin nx,\ \ x \in R$ is uniformly bounded. Infact,

$|f_n(x)| = |\sin nx| \le 1$ for all $x \in R$ and for all $n \in N$.

If each individual function is bounded and if $f_n \to f$ uniformly on E, then it can be shown that $\{f_n\}$ is uniformly bounded on E. This result generally helps us to conclude that a sequence is not uniformly convergent.

2.4 Cauchy Criterion for Uniform Convergence

We now find necessary and sufficient condition for uniform convergence of a sequence of functions.

Theorem 1 (Cauchy criterion for uniform convergence). The sequence of functions $\{f_n\}$, defined on E, converges uniformly if and only if for every $\varepsilon > 0$ there exists an integer N such that $m \ge N, n \ge N, x \in E$ imply $|f_n(x) - f_m(x)| < \varepsilon.$

Proof. Suppose first that $\langle f_n \rangle$ converges uniformly on E to f. Then to each $\varepsilon > 0$ there exists an integer N such that n > N implies

$$|f_n(x) - f(x)| < \frac{\varepsilon}{2}, \quad \text{for all } x \in E$$

Similarly for m > N implies

$$|f_m(x) - f(x)| < \frac{\varepsilon}{2}, \quad \text{for all } x \in E$$

Hence, for n > N, m > N, we have

$$\begin{aligned} |f_n(x) - f_m(x)| &= |f_n(x) - f(x) + f(x) - f_m(x)| \\ &\le |f_n(x) - f(x)| + |f_m(x) + f(x)| \\ &< \varepsilon/2 + \varepsilon/2 = \varepsilon \ \text{ for all } x \in E \end{aligned}$$

Hence the condition is necessary.

Conversely, suppose that the given condition holds. Therefore $\{f_n(x)\}$ is a Cauchy sequence in R for each $x \in E$. Since R is complete, it follows that $\{f_n(x)\}$ converges to some value f(x), for each $x \in E$ & $\{f_n\}$ converges to f pointwise. We need only to show that the convergence is uniform. to show this let $\varepsilon > 0$ be given, then by hypothesis, $n_0 \in N$ (depending only on ε) such that

$$|f_n(x) - f_m(x)| < \varepsilon, \qquad n, m > N \text{ and } x \in E$$

Let n be fixed and let $m \to \infty$, then we have

$$|f_n(x) - f(x)| < \varepsilon \quad \forall \ \ x \in E$$

Hence $f_n \to f$ uniformly on E.

We now find necessary and sufficient condition for uniform convergence of a series of functions.

Theorem 2 (Cauchy criterion for uniform convergence). A series of real functions $\sum f_n$, each defined on a set X converges uniformly on X iff for every $\varepsilon > 0, \exists\ n_0 \in N$ (depending only on ε) such that

$$\left|f_{n+1}(x) + f_{n+2}(x) + ... + f_{n+m}(x)\right| < \varepsilon \qquad \text{for } n \geq n_0, m \geq 1, x \in X.$$

Proof. Let $S_n(x) = f_1(x) + f_2(x) + ... + f_n(x), \forall x \in X$ be a partial sum

$$\sum_{i=1}^{n} f_i(x) = f_1(x) + f_2(x) + ... + f_n(x), \quad x \in X$$

so that $\{S_n(x)\}$ is a sequence of partial sums of the series $\sum_{n=1}^{\infty} f_n$. Now the series $\sum f_n$ is uniformly convergent iff the sequence $\{S_n\}$ is uniformly convergent.

i.e., for given $\varepsilon > 0, \exists$ a positive integer m such that $n \geq m$

$$\left|S_{n+m}(x) - S_n(x)\right| < \varepsilon, m = 1,2,.......... \quad \text{[By Cauchy criteria of uniform converge of sequence]}$$

$$\Leftrightarrow \left|f_{n+1}(x) + f_{n+2}(x) + ... + f_{n+m}(x)\right| < \varepsilon, m = 1,2,........$$

This completes the proof of Cauchy's Criteria for Series.

2.5 Tests for Uniform Convergence

In this section, we study M_n-test, Weierstrass M-test, Abel's Test and Dirichlet's Test for uniform convergence and some examples which emphasis on the applications of these tests.

Theorem 1. Suppose $\lim_{n\to\infty} f_n(x) = f(x),\ x \in E$ and let $M_n = \operatorname{lub}_{x\in E}\left|f_n(x) - f(x)\right|$. Then $f_n \to f$ uniformly on E if and only if $M_n \to 0$ as $n \to \infty$. (This result is known as M_n **- Test for uniform convergence**)

Proof. We have

$$\operatorname{lub}_{x\in E}\left|f_n(x) - f(x)\right| = M_n \to 0 \text{ as } n \to \infty.$$

Hence $$\lim_{n\to\infty}\left|f_n(x) - f(x)\right| = 0 \text{ for all } x \in E.$$

Hence to each $\varepsilon > 0$, there exists an integer N such that $n > N,\ x \in E$ imply

$$\left|f_n(x) - f(x)\right| < \varepsilon$$

Hence $f_n \to f$ uniformly on E.

Example 1. By using M_n – test, show that the sequences $\{f_n\}$ where

$f_n(x) = \dfrac{nx}{1+n^2x^2}$ is not uniformly convergent on any interval containing 0.

Solution. Here $f(x) = \lim_{n\to\infty} f_n(x) = \lim_{n\to\infty} \dfrac{nx}{1+n^2x^2}$

$$= \lim_{n\to\infty} \frac{x/n}{1/n^2 + x^2} = 0$$

Thus the sequence $\{f_n\}$ converges pointwise to the function f identically 0.

Now $M_n = \sup_{x\in[a,b]} |f_n(x) - f(x)|$

$$= \sup_{x\in[a,b]} \left| \frac{nx}{1+n^2x^2} - 0 \right| = \sup_{x\in[a,b]} \left| \frac{nx}{1+n^2x^2} \right|$$

Let us find the maximum value of $\dfrac{nx}{1+n^2x^2}$ by second derivative test.

Let $\phi(x) = \dfrac{nx}{1+n^2x^2}$

$$\phi'(x) = \frac{(1+n^2x^2)n - nx.2n^2x}{(1+n^2x^2)^2}$$

Put $\phi'(x) = 0$.

Then we have, $(1+n^2x^2)n - nx(2n^2x) = 0$

$$(1+n^2x^2) = x(2xn^2)$$

$$1 = 2x^2n^2 - n^2x^2 \Rightarrow 1 = n^2x^2$$

$$\Rightarrow x^2 = \frac{1}{n^2} \Rightarrow x = \pm\frac{1}{n}.$$

or $\quad x = \dfrac{1}{n} \, or - \dfrac{1}{n}.$

Also,

$$\varphi''(x) = \frac{(1+n^2x^2)^2 .n(-2n^2x) - n(1-n^2x^2)(4n^2x + 4n^4x^3)}{(1+n^2x^2)^4}$$

$$= \frac{-2n^3x\left(1+n^2x^2\right)^2 - n\left(1-n^2x^2\right)\left(4n^2x+4n^4x^3\right)}{\left(1+n^2x^2\right)^4}.$$

At $x = \frac{1}{n}$, $\varphi''(x) < 0$. Therefore d(x) is maximum when $x = \frac{1}{n}$.

Also $\phi\left(\frac{1}{n}\right) = \frac{1}{2}$

Thus we take an interval [a,b] containing zero ,then

$$M_n = \sup_{x\in[a,b]} \left|f_n(x) - f(x)\right| = \sup_{x\in[a,b]} \left|\frac{nx}{1+n^2x^2}\right| = \frac{1}{2},$$

which does not tend to zero as $n \to \infty$.

Hence by M_n – test the sequence $\{f_n\}$ is not uniformly continuous in any interval containing zero.

Example 2. Show that the sequence $\{f_n\}$, where

$f_n(x) = \frac{x}{1+nx^2}$ converges uniformly on R.

Solution. Here pointwise limit is

$$f(x) = \lim_{n\to\infty} \frac{x}{1+nx^2} = 0 \ \forall \ x \in R.$$

Let $\phi(x) = f_n(x) - f(x) = \frac{x}{1+nx^2}$.

For maximum & minimum of $\phi(x)$, we have

$$\phi'(x) = 0 \Rightarrow \frac{1+nx^2-2nx^2}{\left(1+nx^2\right)^2} = 0$$

$$\Rightarrow \frac{1-nx^2}{\left(1+nx^2\right)^2} = 0 \Rightarrow 1-nx^2 = 0 \Rightarrow x = \pm\frac{1}{\sqrt{n}}.$$

Now, $\varphi'(x) = \frac{1-nx^2}{\left(1+nx^2\right)^2} = \frac{-\left(1+nx^2\right)}{\left(1+nx^2\right)^2} + \frac{2}{\left(1+nx^2\right)^2}.$

$$= -\frac{1}{\left(1+nx^2\right)} + \frac{2}{\left(1+nx^2\right)^2}.$$

$$\varphi''(x) = \frac{2nx}{\left(1+nx^2\right)^2} - \frac{8nx}{\left(1+nx^2\right)^3}..$$

Put $x = \frac{1}{\sqrt{n}}$,

$$\therefore \quad \varphi''\left(\frac{1}{\sqrt{n}}\right) = \frac{2\sqrt{n}}{2^2} - \frac{8\sqrt{n}}{2^3} = \frac{\sqrt{n}}{2} - \sqrt{n} = -\frac{\sqrt{n}}{2}$$

$\phi''\left(\frac{1}{\sqrt{n}}\right) = -ve$ maximum.

Hence max. $\varphi(x) = \frac{1/\sqrt{n}}{1+n(1/n)} = \frac{1}{2\sqrt{n}}$.

Thus $M_n = \sup\limits_{x\in R}\left|f_n(x) - f(x)\right|$

$$= \sup_{x\in R}\left|\frac{x}{1+nx^2} - 0\right| = \sup_{x\in R}\left|\phi(x)\right| = \frac{1}{2\sqrt{n}}$$

Also so $\lim\limits_{n\to\infty} M_n = \lim\limits_{n\to\infty} \frac{1}{2\sqrt{n}} = 0$

Hence by $M_n -$ test, the sequence $\{f_n(x)\}$ uniformly converges on R.

Example 3. Show that 0 is a point of non – uniformly convergent of the sequence $\{f_n(x)\}$, where $f_n(x) = nxe^{-nx}; x \geq 0$.

Solution. Here pointwise limit,

$$f(x) = \lim_{n\to\infty} f_n(x) = \lim_{n\to\infty} nxe^{-nx} \qquad \left[\frac{\infty}{\infty} form\right].$$

By L'Hospital rule, we get

$$= \lim_{n\to\infty} \frac{x}{xe^{nx}} = 0$$

For maximum & minimum value of $\phi(x)$, where

$$\phi(x) = f_n(x) - f(x) = nxe^{-nx}$$

$$\phi'(x) = nx\left(-ne^{-nx}\right) + ne^{-nx}$$

Now $\quad \phi'(x) = 0 \Rightarrow -n^2xe^{-nx} + ne^{-nx} = 0$

$$x = \frac{-ne^{-nx}}{-n^2e^{-nx}} = \frac{1}{n} \Rightarrow x = \frac{1}{n}$$

Now $\quad \phi''(x) = -n^2x\left(-ne^{-nx}\right) + \left(-n^2e^{-nx}\right) + \left(-n^2e^{-nx}\right)$

$$= n^3 xe^{-nx} - 2n^2 e^{-nx}$$

$$\phi''\left(\frac{1}{n}\right) = n^3 . \frac{1}{n} . \frac{1}{e} - \frac{2n^2}{e} = -\frac{n^2}{e}$$

$\therefore \phi''(x) = -ve$ i.e., maximum at $x = \frac{1}{n}$

Hence max. of $\phi(x) = n . \frac{1}{n} e^{-1} = \frac{1}{e}$.

Thus $M_n = \sup_{x \in R} \left| f_n(x) - f(x) \right|$

$$= \sup_{x \in R} \left| nxe^{-nx} - 0 \right| = \sup_{x \in R} \left| \phi(x) \right| = \frac{1}{e}$$

So $\lim_{n \to \infty} M_n = \lim_{n \to \infty} \frac{1}{e} = \frac{1}{e} \neq 0$

Hence by $M_n -$ test, the sequence of function is not uniform convergent on R.

Weierstrass contributed a very convenient test for the uniformly convergence of infinite series of functions.

Theorem 2 (Weierstrass M-test). Let $\langle f_n \rangle$ be a sequence of functions defined on E and suppose $\left| f_n(x) \right| \le M_n$ ($x \in E$, n = 1, 2, 3,……), where M_n is independent of x. Then $\sum f_n$ converges uniformly as well as absolutely on E if $\sum M_n$ converges.

Proof. Absolute convergence follows immediately from comparison test.

To prove uniform convergence, we note that

$$\left| S_m(x) - S_n(x) \right| = \left| \sum_{i=1}^{m} f_n(x) - \sum_{i=1}^{m} f_n \right|.$$

$$= \left| f_{n+1}(x) + f_{n+2}(x) + ... + f_m(x) \right|$$

$$\le M_{n+1} + M_{n+2} + ... + M_m .$$

But since $\sum M_n$ is convergent, given $\varepsilon > 0$, there exists N (independent of x) such that

$$\left| M_{n+1} + M_{n+2} + ... + M_m \right| < \varepsilon, \ n > N.$$

Hence

$$\left| S_m(x) - S_n(x) \right| < \varepsilon, \ n > N, \ x \in E$$

and so $\sum f_n(x)$ converges uniformly by Cauchy criterion for uniform convergence.

Example 4. Consider the series $\sum_{n=1}^{\infty} \frac{\cos n\theta}{n^p}$. We observe that

$$\left| \frac{\cos n\theta}{n^p} \right| \leq \frac{1}{n^p}.$$

Also, we know that $\sum_{n=1}^{\infty} \frac{1}{n^p}$ is convergent if p > 1. Hence, by Weierstrass M-Test, the series $\sum \frac{\cos n\theta}{n^p}$ converges absolutely and uniformly for all real values of θ if p > 1. Similarly, the series $\sum_{n=1}^{\infty} \frac{\sin n\theta}{n^p}$ converges absolutely and uniformly by Weierstrass's M-Test.

Example 5. Taking $M_n = r^n$, 0 < r < 1, it can be shown by Weierstrass's M-Test that the series $\sum r^n \cos n\theta,\ \sum r^n \sin n\theta,\ \sum r^n \cos^2 n\theta,\ \sum r^n \sin^2 n\theta$ converge uniformly and absolutely.

Example 6. Consider $\sum_{n=1}^{\infty} \frac{x}{n(1+nx^2)},\ x \in R$.

We assume that x is positive, for if x is negative, we can change signs of all the terms. We have

$f_n(x) = \frac{x}{n(1+nx^2)}$ and $f_n'(x) = 0$ implies $nx^2 = 1$. Thus maximum value of $f_n(x)$ is $\frac{1}{2n^{3/2}}$.

Hence
$$f_n(x) \leq \frac{1}{2n^{3/2}}$$

Since $\sum \frac{1}{n^{3/2}}$ is convergent, Weierstrass's M-Test implies that $\sum_{n=1}^{\infty} \frac{x}{n(1+nx^2)}$ is uniformly convergent for all $x \in R$.

Example 7. Consider the series $\sum_{n=1}^{\infty} \frac{x}{(n+x^2)^2},\ x \in R$. We have

$$f_n(x) = \frac{x}{(n+x^2)^2}$$

and so
$$f_n'(x) = \frac{(n+x^2)^2 - 2x(n+x^2)2x}{(n+x^2)^4}$$

Thus $f_n'(x) = 0$ gives

$$x^4 + n^2 + 2nx^2 - 4nx^2 - 4x^4 = 0$$

$$n^2 - 2nx^2 - 3x^4 = 0$$

$$3x^4 + 2nx^2 - n^2 = 0$$

$$x^2 = \frac{n}{3} \text{ or } x = \sqrt{\frac{n}{3}}.$$

Also $f_n''(x)$ is negative. Hence maximum value of $f_n(x)$ is $\frac{3\sqrt{3}}{16n^{3/2}}$. Since $\sum \frac{1}{n^{3/2}}$ is convergent, it follows by Weierstrass's M-Test that the given series is uniformly convergent.

Example 8. The series $\sum_{n=1}^{\infty} \frac{a_n x^n}{1+x^{2n}}$ and $\sum_{n=1}^{\infty} \frac{a_n x^{2n}}{1+x^{2n}}$

converge uniformly for all real values of x and $\sum a_n$ is absolutely convergent. The solution follow the same line as for example 7.

Lemma 1 (Abel's Lemma). If $v_1, v_2, ..., v_n$ be positive and decreasing, the sum $u_1 v_1 + u_2 v_2 + ... + u_n v_n$ lies between $A v_1$ and $B v_1$, where A and B are the greatest and least of the quantities $u_1, u_1 + u_2, u_1 + u_2 + u_3, ..., u_1 + u_2 + ... + u_n$.

Proof. Write

$$S_n = u_1 + u_2 + ... + u_n.$$

Therefore

$$u_1 = S_1, u_2 = S_2 - S_1, ..., u_n = S_n - S_{n-1}.$$

Hence

$$\sum_{i=1}^{n} u_i v_i = u_1 v_1 + u_2 v_2 + ... + u_n v_n.$$

$$= S_1 v_1 + (S_2 - S_1) v_2 + (S_3 - S_2) v_3 + ... + (S_n - S_{n-1}) v_n$$

$$= S_1 (v_1 - v_2) + S_2 (v_2 - v_3) + ... + S_{n-1} (v_{n-1} - v_n) + S_n v_n$$

$$< A(v_1 - v_2 + v_2 - v_3 + ... + v_{n-1} - v_n + v_n).$$

$$= Av_1.$$

Similarly, we can show that

$$\sum_{i=1}^{n} u_i v_i > Bv_1.$$

Hence the result follows.

Theorem 3 (Abel's Test). The series $\sum_{n=1}^{\infty} u_n(x)v_n(x)$ converges uniformly on E if

(i) $\{v_n(x)\}$ is a positive decreasing sequence for all values of $x \in E$

(ii) $\sum u_n(x)$ is uniformly convergent

(iii) $v_1(x)$ is bounded for all $x \in E$, i.e., $v_1(x) < M$.

Proof. Consider the series $\sum u_n(x)v_n(x)$, where $\{v_n(x)\}$ is a positive decreasing sequence for each $x \in E$. By Abel's Lemma

$$|u_n(x)v_n(x) + u_{n+1}(x)v_{n+1}(x) + ... + u_m(x)v_m(x)| < A\, v_n(x),$$

where A is greatest of the magnitudes

$$|u_n(x)|, |u_n(x) + u_{n+1}(x)|, ..., |u_n(x) + u_{n+1}(x) + ... + u_m(x)|.$$

Clearly A is function of x.

Since $\sum u_n(x)$ is uniformly convergent, it follows that

$$|u_n(x) + u_{n+1}(x) + ... + u_m(x)| < \frac{\varepsilon}{M} \text{ for all } n > N,\ x \in E$$

and so $A < \frac{\varepsilon}{M}$ for all n > N (independent of x) and for all $x \in E$. Also, since $\{v_n(x)\}$ is decreasing, $v_n(x) < v_1(x) < M$ since $v_1(x)$ is bounded for all $x \in E$

Hence

$$|u_n(x)v_n(x) + u_{n+1}(x)v_{n+1}(x) + ... + u_m(x)v_m(x)| < \varepsilon$$

for $n > N$ and all $x \in E$ and so $\sum_{n=1}^{\infty} u_n(x)v_n(x)$ is uniformly convergent.

Example 9. Consider the series

$$\sum_{n=1}^{\infty} \frac{(-1)^n}{n^p} \frac{x^{2n}}{1+x^{2n}}.$$

We note that if p > 1, then $\sum \frac{(-1)^n}{n^p}$ is absolutely convergent and is independent of x. Hence, by Weierstrass's M-Test, the given series is uniformly convergent for all $x \in R$.

If $0 \le p \le 1$, the series $\sum \frac{(-1)^n}{n^p}$ is convergent but not absolutely. Let

$$v_n(x) = \frac{x^{2n}}{1+x^{2n}}$$

Then $\langle v_n(x) \rangle$ is monotonically decreasing sequence for $|x| < 1$, because

$$v_n(x) - v_{n+1}(x) = \frac{x^{2n}}{1+x^{2n}} - \frac{x^{2n+2}}{1+x^{2n+2}}$$

$$= \frac{x^{2n}(1-x^2)}{(1+x^{2n})(1+x^{2n+2})} \qquad (+\text{ve})$$

Also $$v_1(x) = \frac{x^2}{1+x^2} < 1.$$

Hence, by Abel's Test, the series $\sum_{n=1}^{\infty} \frac{(-1)^n}{n^p} \cdot \frac{x^{2n}}{1+x^{2n}}$ is uniformly convergent for $0 < p \le 1$ and $|x| < 1$.

Example 10. Consider the series $\sum a_n \cdot \frac{x^n}{1+x^{2n}}$, under the condition that $\sum a_n$ is convergent. Let

$$v_n(x) = \frac{x^n}{1+x^{2n}}$$

Then

$$\frac{v_n(x)}{v_{n+1}(x)} = \frac{1+x^{2n+2}}{x(1+x^{2n})}$$

and so

$$\frac{v_n(x)}{v_{n+1}(x)} - 1 = \frac{(1-x)(1-x^{2n+1})}{x(1+x^{2n})}$$

which is positive if $0 < x < 1$. Hence $v_n > v_{n+1}$ and so $\langle v_n(x) \rangle$ is monotonically decreasing and positive. Also $v_1(x) = \frac{x}{1+x^2}$ is bounded. Hence, by Abel's test, the series $\sum a_n \cdot \frac{x^n}{1+x^{2n}}$ is uniformly convergent in (0, 1) if $\sum a_n$ is convergent.

Example 11. Consider the series $\sum a_n \frac{nx^{n-1}(1-x)}{1-x^n}$ under the condition that $\sum a_n$ is convergent. We have

$$v_n(x) = \frac{nx^{n-1}(1-x)}{1-x^n}.$$

Then

$$\frac{v_n(x)}{v_{n+1}(x)} = \frac{n}{(n+1)x} \cdot \frac{1-x^{n+1}}{(1-x^n)}.$$

Since $\frac{n}{(n+1)} \to 0$ as $n \to \infty$, taking n sufficient large

$$\frac{v_n(x)}{v_{n+1}(x)} > \frac{(1-x^{n+1})}{(1-x^n)} > 1 \text{ if } 0 < x < 1.$$

Hence $<v_n(x)>$ is monotonically decreasing and positive. Hence, by Abel's Test, the given series converges uniformly in (0, 1).

Theorem 4. (Dirichlet's Test for uniform convergence). The series $\sum_{n=1}^{\infty} u_n(x)v_n(x)$ converges uniformly on E if

(i) $\{v_n(x)\}$ is a positive decreasing sequence for all values of $x \in E$, which tends to zero uniformly on E

(ii) $\sum u_n(x)$ oscillates or converges in such a way that the moduli of its limits of oscillation remains less than a fixed number M for all $x \in E$.

Proof. Consider the series $\sum_{n=1}^{\infty} u_n(x)v_n(x)$ where $\{v_n(x)\}$ is a positive decreasing sequence tending to zero uniformly on E. By Abel's Lemma

$$|u_n(x)v_n(x) + u_{n+1}(x)v_{n+1}(x) + ... + u_m(x)v_m(x)| < Av_n(x),$$

where A is greatest of the magnitudes

$$|u_n(x)|, |u_n(x) + u_{n+1}(x)|, ..., |u_n(x) + u_{n+1}(x) + ... + u_m(x)|$$

and A is a function of x.

Since $\sum u_n(x)$ converges or oscillates finitely in such a way that $\left|\sum_{r}^{s} u_n(x)\right| < M$ for all $x \in E$,

therefore A is less than M. Furthermore, since $v_n(x) \to 0$ uniformly as $n \to \infty$, to each $\varepsilon > 0$ there exists an integer N such that

$$v_n(x) < \frac{\varepsilon}{M} \text{ for all } n > N \text{ and all } x \in E.$$

Hence

$$|u_n(x)v_n(x) + u_{n+1}(x)v_{n+1}(x) + ... + u_m(x)v_m(x)| < \frac{\varepsilon}{M}. M = \varepsilon$$

for all $n > N$ and $x \in E$ and so $\sum_{n=1}^{\infty} u_n(x)v_n(x)$ is uniformly convergent on E.

<u>Another way of Dirichlet's Test for uniform convergence with proof.</u>

Statement. If $\{V_n(x)\}$ is a monotonic function of x for each fixed value of x in $[a,b]$ and $\{V_n(x)\}$ converges uniformly to zero for $a \le x \le b$ and if there is a number M > 0 s.t.

$\left|\sum_{r=1}^{n} U_r(x)\right| \le M \forall n \,\&\, x \in [a,b]$, then the series $\sum V_n(x)U_n(x)$ is uniformly convergent on [a,b].

Proof. Since $\{V_n(x)\}$ converges uniformly to zero thus for any $\varepsilon > 0, \exists$ an integer N (Independent of x) s.t. for all $x \in [a,b]$

$$|V_n(x)| < \frac{\varepsilon}{4M} \;\forall\; n \ge N \ldots\ldots\ldots\ldots\ldots\ldots\ldots(1).$$

Let $S_n = \sum_{r=1}^{n} U_r(x) \forall n \qquad \& \, x \in [a,b]$

so that $\qquad |S_n(x)| \le M \;\forall\, n \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(2).$

Now consider $\sum_{r=n+1}^{n+p} V_r(x)U_r(x) = V_{n+1}(x)U_{n+1}(x) + \ldots\ldots\ldots\ldots + V_{n+p}(x)U_{n+p}(x)$

$$= V_{n+1}(x)[S_{n+1} - S_n] + V_{n+2}(x)[S_{n+2} - S_{n+1}] + \ldots\ldots\ldots + V_{n+p}(x)[S_{n+p} - S_{n+p-1}]$$

$$= -V_{n+1}(x)S_n + [V_{n+1}(x) - V_{n+2}(x)]S_{n+1} + \ldots\ldots\ldots\ldots + [V_{n+p-1}(x) - V_{n+p}(x)]S_{n+p-1} + V_{n+p}(x)S_{n+p}$$

$$= \sum_{r=n+1}^{n+p-1} [V_r(x) - V_{r+1}(x)]S_r(x) - V_{n+1}(x)S_n(x) + V_{n+p}(x)S_{n+p}(x)$$

$$\Rightarrow \left|\sum_{r=n+1}^{n+p} V_r(x)U_r(x)\right| \le \sum_{r=n+1}^{n+p-1} |V_r(x) - V_{r+1}(x)||S_r(x)| + |V_{n+1}(x)||S_n(x)| + |V_{n+p}(x)||S_{n+p}(x)|$$

$$\le \sum_{r=n+1}^{n+p-1} |V_r(x) - V_{r+1}(x)| M + \frac{\varepsilon}{4M}.M + \frac{\varepsilon}{4M}.M \;\forall\; n \ge N \qquad \textit{(By (1)\&(2))}$$

$$= M|V_{n+1}(x) - V_{n+p}(x)| + \frac{\varepsilon}{2}$$

$$\le M\left[|V_{n+1}(x)| + |V_{n+p}(x)|\right] + \frac{\varepsilon}{2}$$

$$\le M\left[\frac{\varepsilon}{4M} + \frac{\varepsilon}{4M}\right] + \frac{\varepsilon}{2} = \varepsilon.$$

Hence by Cauchy Criteria, the series $\sum_{r=n+1}^{n+p} V_r(x)U_r(x)$ converges uniformly on [a,b].

Remark 1. The statement $\left|\sum_{r=1}^{n} U_r(x)\right| \le K \ \forall\, x \in [a,b]$ & $\forall\, n$ is equivalent to saying that the sequence of partial sum of series $\sum U_n(x)$ is bounded for each value of $x \in [a,b]$ i.e, for every point $x_i \in [a,b]$, there is a number k_i such that $\left|\sum_{r=1}^{n} U_r(x_i)\right| \le k_i$ and there exists a number k such that $k_i < k \ \forall\ i$.

This fact is also stated as the partial sum of the series is uniformly bounded.

This, in turm is equivalent to saying that the series $\sum u_n(x)$ either converges uniformly or oscillates finitely.

So Dirichlet's test can be states also as "If $V_n(x)$ is a monotonic function of n for each fixed value of x in [a,b] and $V_n(x)$ converges uniformly to zero for $x \in [a,b]$ and if $\sum U_n(x)$ either uniformly converges to zero or oscillates finitely in [a,b]. Then the series $\sum V_n(x)U_n(x)$ is uniformly convergent on [a,b].

Example 12. Prove that the series $\sum \frac{Cosn\phi}{n^p}$ and $\sum \frac{Sinn\phi}{n^p}$ converges uniformly for all values of p > 0 in an interval $[\alpha, 2\pi - \alpha]$ for $0 < \alpha < \pi$.

Solution. When, p > 1 , By Weierstrass M-test at once prove both the series uniformly converge for all values of ϕ.

When $0 < p \le 1$, $\quad U_r = Cosr\phi$

Take $b_n = \frac{1}{n^p}$ and $U_n = Cosn\phi$ or $(Sinn\phi)$

Then by Dirichlet's test $\frac{1}{n^p}$ is positive and monotonic decreasing and uniformly tending to zero with

$$\left|\sum_{r=1}^{n} U_r\right| = \left|\sum_{r=1}^{n} Cosn\phi\right| = \left|Cos\phi + Cos2\phi + \ldots\ldots\ldots\ldots. + Cosn\phi\right|$$

$$= \left|\frac{Sinn\,\phi/2}{Sin\,\phi/2} . Cos\frac{(Ist.angle + Last.angle)}{2}\right|$$

$$= \left|\frac{Sinn\,\phi/2}{Sin\,\phi/2} . Cos\frac{(\phi + n\phi)}{2}\right|$$

$$\le Co\sec\,\phi/2 \ \forall n \qquad (\because\ |\mathrm{Sin}\, n\varphi \le 1|\ and\ |\mathrm{Cos}\,\varphi \le 1|).$$

Thus all the conditions of Dirichlet's test are fulfilled and the series $\sum \frac{Cosn\phi}{n^p}$ and $\sum \frac{Sinn\phi}{n^p}$ converges on $[\alpha, 2\pi - \alpha]$.

2.6 Uniform Convergence and Continuity

We know that if f and g are continuous functions, then f + g is also continuous and this result holds for the sum of finite number of functions. The question arises "Is the sum of infinite number of continuous function a continuous function?". The answer is not necessary. The aim of this section is to obtain sufficient condition for the sum function of an infinite series of continuous functions to be continuous.

Theorem 1. Let $\langle f_n \rangle$ be a sequence of continuous functions on a set $E \subseteq R$ and suppose that $\langle f_n \rangle$ converges uniformly on E to a function $f : E \to R$. Then the limit function f is continuous.

Proof. Let $c \in E$ be an arbitrary point. If c is an isolated point of E, then f is automatically continuous at c. So suppose that c is an accumulation point of E. We shall show that f is continuous at c. Since $f_n \to f$ uniformly, for every $\varepsilon > 0$ there is an integer N such that $n \geq N$ implies

$$\left|f_n(x) - f(x)\right| < \frac{\varepsilon}{3} \text{ for all } x \in E.$$

Since f_M is continuous at c, there is a neighbourhood $S_\delta(c)$ such that $x \in S_\delta(c) \cap E$ (since c is limit point) implies

$$\left|f_M(x) - f_M(c)\right| < \frac{\varepsilon}{3}.$$

By triangle inequality, we have

$$\begin{aligned}\left|f(x) - f(c)\right| &= \left|f(x) - f_M(x) + f_M(x) - f_M(c) + f_M(c) - f(c)\right| \\ &\leq \left|f(x) - f_M(x)\right| + \left|f_M(x) - f_M(c)\right| + \left|f_M(c) - f(c)\right| \\ &< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon\end{aligned}$$

Hence

$$\left|f(x) - f(c)\right| < \varepsilon, \; x \in S_\delta(\varepsilon) \cap E$$

which proves the continuity of f at arbitrary point $c \in E$.

Remark 1. Uniform convergence of $\langle f_n \rangle$ in above theorem is sufficient but not necessary to transmit continuity from the individual terms to the limit function. For example, let $f_n : [0,1] \to R$ be defined for $n \geq 2$ by

$$f_n(x) = \begin{cases} n^2 x & \text{for } 0 \le x \le \dfrac{1}{n} \\ -n^2\left(x - \dfrac{2}{n}\right) & \text{for } \dfrac{1}{n} \le x \le \dfrac{2}{n}. \\ 0 & \text{for } \dfrac{2}{n} \le x \le 1 \end{cases}$$

Each of the function f_n is continuous on [0, 1]. Also $f_n(x) \to 0$ as $n \to \infty$ for all $x \in [0,1]$. Hence the limit function f vanishes identically and is continuous. But the convergence $f_n \to f$ is non-uniform.

The series version of Theorem 1 is the following:

Theorem 2. If the series $\sum f_n(x)$ of continuous functions is uniformly convergent to a function f on [a, b], then the sum function f is also continuous on [a, b].

Proof. Let $S_n(x) = \sum_{i=1}^{n} f_i(x), n \in N$ and let $\varepsilon > 0$. Since $\sum f_n$ converges uniformly to f on [a, b], there exists a positive integer N such that

$$|S_n(x) - f(x)| < \frac{\varepsilon}{3} \quad \text{for all } n \ge N \text{ and } x \in [a,b] \qquad (1).$$

Let c be any point of [a, b], then (1) implies

$$|S_n(c) - f(c)| < \frac{\varepsilon}{3} \quad \text{for all } n \ge N \qquad (2).$$

Since f_n is continuous on [a, b] for each n, the partial sum

$$S_n(x) = f_1(x) + f_2(x) + ... + f_n(x)$$

is also continuous on [a, b] for all n. Hence to each $\varepsilon > 0$ there exists a $\delta > 0$ such that

$$|S_n(x) - S_n(c)| < \frac{\varepsilon}{3} \text{ whenever } |x - c| < \delta \qquad (3).$$

Now, by triangle inequality, and using (1), (2) and (3), we have

$$\begin{aligned} |f(x) - f(c)| &= |f(x) - S_n(x) + S_n(x) - S_n(c) + S_n(c) - f(c)| \\ &\le |f(x) - S_n(x)| + |S_n(x) - S_n(c)| + |S_n(c) - f(c)| \\ &< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon, \text{ whenever } |x - c| < \delta. \end{aligned}$$

Hence f is continuous at c. Since c is arbitrary point in [a, b], f is continuous on [a, b].

However, the converse of Theorem 1 is true with some additional condition on the sequence $\langle f_n \rangle$ of continuous functions. The required result goes as follows:

Theorem 3. (Dini‘s theorem on uniform convergence of subsequences (first form)). Let E be compact and let $\{f_n\}$ be a sequence of functions continuous on E which converges to a continuous function f on E. If $f_n(x) \geq f_{n+1}(x)$ for n = 1, 2, 3, …, and for every $x \in E$, then $f_n \to f$ uniformly on E.

Proof. Take

$$g_n(x) = f_n(x) - f(x).$$

Being the difference of two continuous functions $g_n(x)$ is continuous. Also $g_n \to 0$ and $g_n \geq g_{n+1}$. We shall show that $g_n \to 0$ uniformly on E.

Let $\varepsilon > 0$ be given. Since $g_n \to 0$, there exists an integer $n \geq N_x$ such that

$$|g_n(x) - 0| < \varepsilon / 2$$

In particular

$$|g_{N_x}(x) - 0| < \varepsilon / 2$$

i.e. $$0 \leq g_{N_x}(x) < \varepsilon / 2.$$

The continuity and monotonicity of the sequence $\{g_n\}$ imply that there exists an open set J(x) containing x such that

$$0 \leq g_n(t) < \varepsilon$$

if $t \in J(x)$ and $n \geq N_x$.

Since E is compact, there exists a finite set of points $x_1, x_2, \ldots., x_m$ such that

$$E \subseteq J(x_1) \cup J(x_2) \cup \ldots \cup J(x_m).$$

Taking

$$N = \max\{N_{x_1}, N_{x_2}, \ldots, N_{x_m}\}.$$

it follows that

$$0 \leq g_n(t) \leq \varepsilon$$

for all $t \in E$ and $n \geq N$. Hence $g_n \to 0$ uniformly on E and so $f_n \to f$ uniformly on E.

Theorem 4. If a sequence $\{f_n\}$ of real valued function converges uniformly to f in [a,b] and let x_0 be a point of [a,b] s.t. $\lim_{x \to x_0} f_n(x) = a_n; \quad (n = 1, 2, \ldots\ldots\ldots).$.

Then (i) $\{a_n\}$ converges.

(ii) $\lim_{x \to x_0} f(x) = \lim_{n \to \infty} a_n.$

i.e, $\lim\limits_{x\to x_0}\lim\limits_{n\to\infty} f_n(x) = \lim\limits_{n\to\infty}\lim\limits_{x\to x_0} f_n(x)$.

Proof. (i) The sequence $\{f_n\}$ converges uniformly on [a,b] .Therefore for $\varepsilon > 0$, there exists an integer m (independent of x) s.t. for all $x \in [a,b]$

$$\left|f_{n+p}(x) - f(x)\right| < \varepsilon \ \forall \ n > m, p \geq 1 \quad \text{(By Cauchy's Criterion).}$$

Keeping n, p fixed and tending $x \to x_0$, we get

$$\left|a_{n+p} - a_n\right| < \varepsilon \ \forall n \geq m, p \geq 1$$

So that $\{a_n\}$ is a Cauchy sequence and therefore converges to A.

(ii) Since $\{f_n\}$ converges uniformly to f.

Thus for given $\varepsilon > 0$, there exists an integer N_1 s.t. for all $x \in [a,b]$.

$$\left|f_n(x) - f(x)\right| < \frac{\varepsilon}{3} \quad \forall \quad n \geq N_1 \qquad (1)$$

Now the sequence $\{a_n\}$ converges to A. So there exists an integer N_2 s.t.

$$\left|a_n - A\right| < \frac{\varepsilon}{3} \quad \forall \quad n \geq N_2 \qquad (2)$$

Now take a no. N such that $N = \max.\{N_1, N_2\}$

Since we have ,

$$\lim_{x\to x_0} f_n(x) = a_n$$

In particular, $\lim\limits_{x\to x_0} f_N(x) = a_N$

$\Rightarrow$ for $\varepsilon > 0, \exists$ a $\delta > 0$ such that

$$\left|f_N(x) - a_N\right| < \varepsilon/3 \quad \text{whenever } \left|x - x_0\right| < \delta \qquad (3)$$

Now, $\left|f(x) - A\right| = \left|f(x) + f_N(x) - f_N(x) - a_N + a_N - A\right|$

$$\leq \left|f(x) - f_N(x)\right| + \left|f_N(x) - a_N\right| + \left|a_N - A\right|$$

$$< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon \quad \text{whenever } \left|x - x_0\right| < \delta$$

$\Rightarrow \lim\limits_{x\to x_0} f(x)$ exists and is equal to A.

Thus $\lim\limits_{x\to x_0} f(x) = \lim\limits_{n\to\infty} a_n = A$.

Hence the Proof.

Theorem 5. If a series $\sum_{n=1}^{\infty} f_n$ converges uniformly to f in [a,b] and x_0 is a point of [a,b] such that $\lim_{x \to x_0} f_n(x) = a_n; (n = 1,2,...............)$

Then (i) $\sum_{n=1}^{\infty} a_n$ converges

(ii) $\lim_{x \to x_0} f(x) = \sum_{n=1}^{\infty} a_n$

Proof. (i) Given that the series $\sum f_n$ converges uniformly on [a,b], for given $\varepsilon > 0$, there exists an integer m such that for all $x \in [a,b]$

$$\left| \sum_{r=n+1}^{n+p} f_r(x) \right| < \varepsilon \ \forall \ n \geq m, p \geq 1$$

(By Cauchy's Criterion)

Keeping n,p fixed and taking the limits $x \to x_0$, we obtain

$$\left| \sum_{r=n+1}^{n+p} a_r(x) \right| < \varepsilon$$

$\Rightarrow$ the series $\sum a_n$ converges to A.

(ii) Since the series $\sum_{n=1}^{\infty} f_n$ converges uniformly to f, therefore for $\varepsilon > 0$, there exists an integer N_1 such that $\forall x \in [a,b]$, we have,

$$\left| \sum_{r=1}^{n} f_r(x) - f(x) \right| < \frac{\varepsilon}{3} \ \forall \ n \geq N_1 ....(1)$$

Again $\sum a_n$ converges to A.

$\Rightarrow$ for $\varepsilon > 0$, $\exists N_2$ such that

$$\left| \sum_{r=1}^{n} a_r - A \right| < \frac{\varepsilon}{3} \forall n \geq N_2 ....(2)$$

Also it is given that

$$\lim_{x \to x_0} f_n(x) = a_n; (n = 1,2,..........)$$

$\Rightarrow$ for the given $\varepsilon > 0, \exists$ a $\delta_i > 0$ such that for i = 1,2,……..

Such that $\left| f_n(x) - a_n \right| < \frac{\varepsilon}{3\mathrm{N}}$ whenever $\left| x - x_0 \right| < \delta_i$.

If we take $\delta = \min.\{\delta_1, \delta_2, ...\delta_N\}$, then we have

$$\left| f_n(x) - a_n \right| < \frac{\varepsilon}{3N} \quad \text{for} \quad \left| x - x_0 \right| < \delta$$

Thus $\left| \sum_{r=1}^{N} f_r(x) - \sum_{r=1}^{N} a_r \right| \leq \sum_{r=1}^{N} \left| f_r(x) - a_r \right| < N.\frac{\varepsilon}{3N} = \frac{\varepsilon}{3}$...........(3)

Now for $\left| x - x_0 \right| < \delta$, we have

$$\left| f(x) - A \right| \leq \left| f(x) - \sum_{r=1}^{N} f_r(x) \right| + \left| \sum_{r=1}^{N} f_r(x) - \sum_{r=1}^{N} a_r \right| + \left| \sum_{r=1}^{N} a_r - A \right|.$$

Using (1),(2) & (3) , we get

$$\left| f(x) - A \right| < \varepsilon$$

$\Rightarrow \lim_{x \to x_0} f(x)$ exists and is equal to A.

We have seen earlier that if sequence $\{f_n\}$ is a sequence of continuous functions which converges pointwise to the function f, then it is not necessary for f to be continuous. However, the concept of uniform convergence is of much importance as the property of continuity transfers to the limit function if the given sequence converges.

Theorem 6. If the sequence of continuous function $\{f_n\}$ is uniformly convergent to a function f on [a,b] then f is continuous on [a,b].

Proof. Let $\varepsilon > 0$ be given.

Now given that sequence $\{f_n\}$ is uniformly convergent to f on [a, b], then there exists a positive integer m such that

$$\left| f_n(x) - f(x) \right| < \frac{\varepsilon}{3} \forall n \geq m \,\&\, \forall x \in [a,b] \qquad (1)$$

Let x_0 be any point of [a, b].

In particular then from (1),

$$\left| f_n(x_0) - f(x_0) \right| < \frac{\varepsilon}{3} \; \forall \; n \geq m \qquad (2)$$

Now f_n is continuous at $x_0 \in [a,b]$ .So, there exists $\delta > 0$ such that

$$\left|f_n(x)-f_n(x_0)\right|<\frac{\varepsilon}{3} \text{ whenever } \left|x-x_0\right|<\delta \qquad (3)$$

Hence for $\left|x-x_0\right|<\delta$, we have

$$\left|f(x)-f(x_0)\right|=\left|f(x)-f_n(x)+f_n(x)+f_n(x_0)-f_n(x_0)-f(x_0)\right|$$

$$\leq\left|f(x)-f_n(x)\right|+\left|f_n(x)-f_n(x_0)\right|+\left|f_n(x_0)-f(x_0)\right|$$

$$<\frac{\varepsilon}{3}+\frac{\varepsilon}{3}+\frac{\varepsilon}{3}=\varepsilon \qquad \text{(from (1), (2)\& (3))}$$

We get $\left|f(x)-f(x_0)\right|<\varepsilon$ whenever $\left|x-x_0\right|<\delta$

Hence f is continuous at $x_0\in[a,b]$

$\Rightarrow$ f is continuous on [a,b].

Theorem 7. If a series $\sum_{n=1}^{\infty}f_n$ of continuous function is uniformly convergent to a function f on [a, b], then the sum function f is also continuous on [a, b].

Proof. Since the series $\sum f_n$ converges uniformly on [a, b] to f on [a, b].

Thus given $\varepsilon>0$, we can choose m such that

$$\textit{for all } x\in[a,b] \qquad \left|\sum_{r=1}^{n}f_r(x)-f(x)\right|<\frac{\varepsilon}{3}\forall n\geq m. \qquad (1)$$

Let x_0 be any point in [a,b], then from (1),we have n = N

$$\left|\sum_{r=1}^{N}f_r(x_0)-f(x_0)\right|<\frac{\varepsilon}{3}\forall n\geq m \qquad (2)$$

Now it is given that each f_n is continuous on [a,b] and in particular at x_0.

Hence $\varepsilon>0$, there exists $\delta>0$ such that

$$\left|\sum_{r=1}^{N}f_r(x)-\sum_{r=1}^{N}f_r(x_0)\right|<\frac{\varepsilon}{3} \text{ whenever } \left|x-x_0\right|<\delta \qquad (3)$$

Hence for $\left|x-x_0\right|<\delta$,

$$\left|f(x)-f(x_0)\right|=\left|f(x)-\sum_{r=1}^{N}f_r(x)+\sum_{r=1}^{N}f_r(x)-\sum_{r=1}^{N}f_r(x_0)+\sum_{r=1}^{N}f_r(x_0)-f(x_0)\right|$$

$$\leq \left|\sum_{r=1}^{N} f_r(x) - f(x)\right| + \left|\sum_{r=1}^{N} f_r(x_0) - f(x_0)\right| + \left|\sum_{r=1}^{N} f_r(x) - \sum_{r=1}^{N} f_r(x_0)\right|$$

Thus from (1), (2) & (3) , we get

$$< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon$$

$$\Rightarrow \left|f(x) - f(x_0)\right| < \varepsilon$$

$\Rightarrow$ f is continuous at x_0 on [a,b]. Since x_0 was chosen arbitrary.

Hence the proof.

Remark 1. (i) Uniform convergence of the sequence $\{f_n\}$ is sufficient but not a necessary condition for the limit function to be continuous. This means that a sequence of continuous functions may have a continuous limit function without uniform convergence.

However the above theorem yields a negative test for uniform convergence of a sequence namely "If the sequence of continuous functions is discontinuous, the sequence cannot be uniformly convergent."

(ii) The same argument hold good in the case of infinite series $\sum_{n=1}^{\infty} f_n$.

The following examples illustrate the same:

(1) The sequence $\{x^n\}$ of continuous functions has a discontinuous limit function f which is given by

$$f(x) = \begin{cases} 0, & if \quad 0 \leq x < 1 \\ 1, & if \quad x = 1 \end{cases} .$$

Then the sequence cannot uniformly convergent on [0, 1].

(2) The sequence $\left\{\frac{nx}{1+n^2x^2}\right\}$ of continuous functions has a continuous limit function but the given sequence is not uniformly convergent.

(3) The sum of the functions of the series $\sum_{n=1}^{\infty}(1-x)x^n$ of the continuous functions.

$$f(x) = \begin{cases} 1, & if \quad x \neq 0 \\ 0, & if \quad x = 0 \end{cases}$$

which is discontinuous on [0,1]. Therefore the series is not uniformly convergent on [0,1].

Note 1. $(1-x)\sum_{n=1}^{\infty} x^n = (1-x)(1+x+x^2+..........)$

$$= (1-x)\left(\frac{1}{1-x}\right) = 1.$$

Some important results

Here we state some results which we shall use in the following theorems & examples:

(1) Every monotonically increasing sequence bounded above converges to the least upper bound (l.u.b.).

(2) Every monotonically decreasing sequence bounded below converges to greatest lower bound (g.l.b).

(3) A real no. ξ is said to be a limit point of a sequence $\{a_n\}$ if given any $\varepsilon > 0$ and a +ve integer m, there exists a +ve integer k > m such that $|a_k - \xi| < \varepsilon$.

(4) Every bounded sequence has a cluster point.

(5) If a seq. $\{a_n\}$ converges to L or diverges to $+\infty$ *or* $-\infty$ then every subsequence of $\{a_n\}$ also converges to L or diverges to $+\infty$ *or* $-\infty$.

(6) Consider the geometric series

$$a + ar + ar^2 + \ldots\ldots\ldots + ar^{n-1} + \ldots\ldots\ldots$$

This series

(i) converges if r < 1.

(ii) diverges to ∞ if $r \geq 1$.

(iii) oscillate finitely if r = -1.

(iv) oscillates infinitely if r < -1.

(7) **Leibnitz's Rule.** The alternative series $\sum_{n=1}^{\infty}(-1)^{n-1} a_n$ is convergent if

(i) $a_{n+1} < a_n \forall n$

(ii) $a_n \to 0$ as $n \to \infty$

(8) For every limit point of a sequence we can form a subsequence converging to limit point. Limit point is also called subsequential limit.

Theorem 8 (Dini's theorem on uniform convergence of subsequences(2nd form)). If a sequence of continuous function $\{f_n\}$ defined on [a,b] is monotonically increasing & converges pointwise to a continuous function f, then the convergence is uniform on [a,b].

Proof. The sequence $\{f_n\}$ is monotonically increasing and converges to f on [a,b].

Therefore, for any $\varepsilon > 0$ and for a point $x \in [a,b]$ there is an integer N s.t.

$$0 \le f(x) - f_n(x) < \varepsilon \;\forall n \ge N \qquad (1)$$

We consider $R_n = f(x) - f_n(x)$; n = 1,2,…………..

Since the sequence $\{f_n\}$ is monotonically increasing. So, the seq. $\{R_n(x)\}$ is monotonically decreasing.

i.e, $R_1(x) \ge R_2(x) \ge R_3(x) \ge .... \ge R_n(x) \qquad (2)$.

Also, the sequence $\{R_n(x)\}$ is bounded below by 0.

Hence the seq. $\{R_n\}$ converges pointwise to 0 on [a,b].

We claim that this convergence is uniform.

Suppose if possible for a fixed $a_0 > 0, \exists$ no integer N which works for all $x \in [a,b]$.

Then for each n = 1,2,3,………, there exists $x_n \in [a,b]$ such that

$$R_n(x_n) \ge a_0 \qquad (3)$$

The seq. $\{x_n\}$ of points belonging to the interval [a,b] is bounded and thus has atleast one limit say $'\xi'$ in [a,b].

Consequently, we can assume that there is a subsequence $\{x_{n_k}\}$ of seq. $\{x_n\}$ converges to $'\xi'$

i.e, $x_{n_k} \to \xi$ as $k \to \infty$.

Now the function,

$R_n(x) = f(x) - f_n(x)$ is continuous being the difference of two continuous functions and thus for every fixed m, we have

$$\lim_{k \to \infty} R_m(x_{n_k}) = R_m(\xi) \qquad \because x_{n_k} \to \xi \text{ as } k \to \infty.$$

Now for every m and any sufficiently large k, we have

$$n_k \ge m, k > m.$$

Since $\{R_m\}$ is a decreasing sequence, we have

$$R_m(x_{n_k}) \ge R_{n_k}(x_{n_k}) \ge a_0 \qquad \text{(from (3))}$$

$$\Rightarrow R_m(x_{n_k}) \ge a_0.$$

But this is contradiction to the fact that sequence $\{R_m\}$ converges pointwise to 0 i.e.,

$$\lim_{n\to\infty} R_m(\xi) = 0$$

Thus the convergence must be uniform and this completes the proof.

Theorem 9 (Dini's theorem on uniform convergence for series). If the sum function of a series $\sum f_n$ with non negative terms defined on an interval [a, b] is continuous on [a,b], then the series is uniformly convergent on the interval [a, b].

Proof. Consider the partial sum of the given series

$$S_n(x) = \sum_{r=1}^{n} f_r(x)$$

Since all the function f_r are non –ve . So, the seq. of partial sum $\{S_n\}$ should be increasing.

Therefore, $S_n(x) \le S_{n+1}(x)\ \forall\ n$

i.e., $\{S_n\}$ is an increasing sequence of continuous functions converges pointwise to a continuous function f. Hence by Theorem 8, the sequence $\{S_n\}$ converges uniformly and the given series is also uniformly convergent.

This completes the proof.

Example 1. Show that the series

$$x^4 + \frac{x^4}{1+x^4} + \frac{x^4}{(1+x^4)^2} + \text{..................}$$

is not uniformly convergent on [a,b].

Solution. The terms of the given series are quotient of two polynomials and hence continuous (Since the polynomials are continuous and quotient of two continuous function is continuous).

Now, Let us find the sum function for the given series. Let, $f(x)$ denotes the sum function of the given series.

If $x \neq 0$ then the series is a geometric series with common ratio $\frac{1}{1+x^4}$ and $\left|\frac{1}{1+x^4}\right| < 1 \forall x \in [0,1]$.

Hence the sum function is given by

$$f(x) = \frac{x^4}{1 - \frac{1}{1+x^4}} = 1 + x^4$$

Thus, $f(x) = \begin{cases} 1+x^4 & if\ \ x \neq 0 \\ 0, & if\ \ x = 0 \end{cases}$

which is discontinuous on 0 and hence on [0,1] . So, the series cannot converge uniformly on [0,1].

Example 2. Show that the series $\sum \frac{x}{(nx+1)\{(n-1)x+1\}}$ is uniformly convergent on any interval [a, b], 0 < a < b, but only pointwise on [0, b].

Solution. Let $f_n(x) = \frac{x}{(nx+1)\{(n-1)x+1\}} = \frac{1}{(n-1)x+1} - \frac{1}{nx+1}$

Therefore n[th] partial sum is

$$S_n(x) = \sum_{r=1}^{n} f_r(x) = f_1(x) + f_2(x) + \dots\dots\dots + f_n(x)$$

$$= \left(1 - \frac{1}{x+1}\right) + \left(\frac{1}{x+1} - \frac{1}{2x+1}\right) + \dots\dots\dots + \left(\frac{1}{(n-1)x+1} - \frac{1}{nx+1}\right)$$

$$= 1 - \frac{1}{nx+1}$$

The sum function $f(x) = \lim_{n\to\infty} S_n(x)$

$$= \lim_{n\to\infty}\left(1 - \frac{1}{nx+1}\right)$$

$$= \begin{cases} 1 & if \quad x \neq 0 \\ 0, & if \quad x = 0 \end{cases}$$

Clearly f is discontinuous at x = 0 and hence discontinuous on [0, b].

This implies that the convergence is not uniform on [0, b] i.e, it is only pointwise.

Now take the interval [a,b] such that 0 < a < b, then the given series is uniformly convergent on [a,b] if for given $\varepsilon > 0$.

$$|S_n(x) - f(x)| = \frac{1}{nx+1} < \varepsilon$$

i.e, if $n > \frac{1}{x}\left(\frac{1}{\varepsilon} - 1\right)$

Now, $\frac{1}{x}\left(\frac{1}{\varepsilon} - 1\right)$ decreasing with x and its maximum value is

$\frac{1}{a}\left(\frac{1}{\varepsilon} - 1\right) = m_0$ (say).

If we take $m > m_0$ then for all $x \in [a,b]$

$$|S_n(x) - f(x)| < \varepsilon \;\forall n \geq m.$$

Hence the series converges uniformly on [a,b] s.t. 0 < a < b.

Example 3. Show that the series $\sum_{n=1}^{\infty}\frac{(-1)^{n-1}}{n+x^2}$ is uniformly convergent but not absolutely for all real values of x.

Solution. The given series is $\sum_{n=1}^{\infty}\frac{(-1)^{n-1}}{n+x^2}$.

Let $a_n=\frac{1}{n+x^2}$.

$a_{n+1}<a_n \forall n$ and $a_n \to 0$ as $n\to\infty$.

Hence by Leibnitz's rule, the alternative series $\sum_{n=1}^{\infty}\frac{(-1)^{n-1}}{n+x^2}$ is convergent.

We know that a series $\sum_{n=1}^{\infty}a_n$ is said to be absolutely convergent if the series $\sum_{n=1}^{\infty}|a_n|$ is convergent.

Now, $\sum_{n=1}^{\infty}\left|\frac{(-1)^{n-1}}{n+x^2}\right|=\sum_{n=1}^{\infty}\frac{1}{n+x^2}$ which behaves like $\sum 1/n$ and hence is divergent.

It remains to prove that the given series is uniformly convergent.

Let $S_n(x)$ denotes the partial sum and $S(x)$ denote the sum of the series.

Now, consider

$$S_{2n}(x)=\frac{1}{1+x^2}-\frac{1}{2+x^2}+\frac{1}{3+x^2}-\frac{1}{4+x^2}+\ldots\ldots\ldots\ldots-\frac{1}{2n+x^2}$$

$$\Rightarrow S_{2n}(x)=\left(\frac{1}{1+x^2}-\frac{1}{2+x^2}\right)+\left(\frac{1}{3+x^2}-\frac{1}{4+x^2}\right)+\ldots\ldots\ldots\ldots\ldots\ldots+\left(\frac{1}{(2n-1)+x^2}-\frac{1}{2n+x^2}\right)$$

Now, note that each bracket in the above expression is positive. Hence $S_{2n}(x)$ is positive and increasing to the sum $S(x)$.

$$\Rightarrow S(x)-S_{2n}(x)>0$$

Also $S(x)-S_{2n}(x)=\frac{1}{(2n+1)+x^2}-\frac{1}{(2n+2)+x^2}+\frac{1}{(2n+3)+x^3}-\ldots\ldots$

$$=\frac{1}{(2n+1)+x^2}-\left(\frac{1}{(2n+2)+x^2}-\frac{1}{(2n+3)+x^2}\right)-\ldots\ldots\ldots\ldots$$

$$< \frac{1}{(2n+1)+x^2}$$

$$< \frac{1}{2n+1}.$$

So, $0 < S(x) - S_{2n}(x) < \frac{1}{2n+1}$ (1).

Also, consider

$$S_{2n+1}(x) - S(x) = \frac{1}{(2n+2)+x^2} - \frac{1}{(2n+3)+x^2} + \frac{1}{(2n+4)+x^2} - \text{.............}$$

$$= \left(\frac{1}{(2n+2)+x^2} - \frac{1}{(2n+3)+x^2}\right) + \left(\frac{1}{(2n+4)+x^2} - \frac{1}{(2n+5)+x^2}\right) + \text{.............}$$

$$S_{2n+1}(x) - S(x) < \frac{1}{(2n+2)+x^2} < \frac{1}{2n+2}.$$

$$\Rightarrow 0 < S_{2n+1}(x) - S(x) < \frac{1}{2n+2} < \frac{1}{2n+1} \qquad (2).$$

Inequality (1) & (2) yield that for any $\varepsilon > 0$,

we can choose an integer m s.t. for all values of x.

$$|S(x) - S_n(x)| < \varepsilon \ \forall\, n \geq m$$

$\Rightarrow$ The series converges uniformly for all real values of x.

Example 4. Consider the seq. $\{f_n\}$ where

$$f_n(x) = \frac{nx}{1+n^2x^2}.$$

Show that the sequence of differentiable functions $\{f_n\}$ does not converge uniformly in an interval containing zero.

Solution. Here $f_n(x) = \frac{nx}{1+n^2x^2}$

$$\Rightarrow f(x) = \lim_{n\to\infty} f_n(x) = 0$$

$\Rightarrow f'(x) = 0$ for all real x

Now, $f_n'(x)=\dfrac{(1+n^2x^2)n-2nx.n^2x}{(1+n^2x^2)^2}=\dfrac{n^3}{n^4}\left(\dfrac{1/n^2+x^2-2x^2}{(1/n^2+x^2)^2}\right)$

Now $\lim\limits_{n\to\infty} f_n'(x)=0$ for $x\neq 0$.

Thus $\lim\limits_{n\to\infty} f_n'(x)=f'(x)$

But at x=0; $f_n'(x)=n$ and $\lim\limits_{n\to\infty} f_n'(0)=\infty$

Thus at $x=0$; $f'(x)\neq \lim\limits_{n\to\infty} f_n'(x)$.

Hence the sequence f_n' does not converges uniformly in an interval that contains zero.

2.7 Uniform Convergence and Integrability.

We know that if f and g are integrable, then $\int(f+g)=\int f+\int g$ and this result holds for the sum of a finite number of functions.

The aim of this section is to find sufficient condition to extend this result to an infinite number of functions.

Theorem 1. Let α be monotonically increasing on [a, b]. Suppose that each term of the sequence $\{f_n\}$ is a real valued function such that $f_n \in R(\alpha)$ on [a, b] for n = 1, 2,.. and suppose $f_n \to f$ uniformly on [a, b]. Then $f \in R(\alpha)$ on [a, b] and

$$\int_a^b f\, d\alpha = \lim_{n\to\infty}\int_a^b f_n\, d\alpha,$$

that is, $$\int_a^b \lim_{n\to\infty} f_n(x)\, d\alpha(x) = \lim_{n\to\infty}\int_a^b f_n(x)\, d\alpha(x)$$

(Thus limit and integral can be interchanged in this case. This property is generally described by saying that **a uniformly convergent sequence can be integrated term by term**).

Proof. Let ε be a positive number. Choose $\eta>0$ such that

$$\eta\left[\alpha(b)-\alpha(a)\right]\leq\frac{\varepsilon}{3} \qquad (1)$$

This is possible since α is monotonically increasing. Since $f_n \to f$ uniformly on [a, b], to each $\eta>0$ there exists an integer n such that

$$\left|f_n(x)-f(x)\right|\leq\eta,\quad x\in[a,b] \qquad (2)$$

Since $f_n \in R(\alpha)$, we choose a partition P of [a, b] such that

$$U(P, f_n, \alpha) - L(P, f_n, \alpha) < \frac{\varepsilon}{3} \qquad (3)$$

The expression (2) implies

$$f_n(x) - \eta \le f(x) \le f_n(x) + \eta$$

Now $f(x) \le f_n(x) + \eta$ implies, by (1) that

$$U(P, f, \alpha) \le U(P, f_n, \alpha) + \frac{\varepsilon}{3} \qquad (4)$$

Similarly, $f(x) \ge f_n(x) - \eta$ implies

$$L(P, f, \alpha) \ge L(P, f_n, \alpha) - \frac{\varepsilon}{3} \qquad (5)$$

Combining (3), (4) and (5), we get

$$U(P, f, \alpha) - L(P, f, \alpha) < \varepsilon$$

Hence $f \in R(\alpha)$ on [a, b].

Further uniform convergence implies that to each $\varepsilon > 0$, there exists an integer N such that $n \ge N$

$$|f_n(x) - f(x)| < \frac{\varepsilon}{[\alpha(b) - \alpha(a)]}, \quad x \in [a, b]$$

Then for $n > N$,

$$\left|\int_a^b f\, d\alpha - \int_a^b f_n\, d\alpha\right| = \left|\int_a^b (f - f_n)\, d\alpha\right| \le \int_a^b |f - f_n|\, d\alpha$$

$$< \frac{\varepsilon}{[\alpha(b) - \alpha(a)]} \int_a^b d\alpha(x)\, dx$$

$$= \frac{\varepsilon[\alpha(b) - \alpha(a)]}{\alpha(b) - \alpha(a)} = \varepsilon .$$

Hence $$\int_a^b f\, d\alpha = \lim_{n\to\infty} \int_a^b f_n\, d\alpha$$

and the result follows.

The series version of Theorem 1 is

Theorem 2. Let $f_n \in R$, n = 1, 2, … If $\sum f_n$ converges uniformly to f on [a, b], then $f \in R$ and

$\int_a^b f(x)\, d\alpha = \sum_{n=1}^{\infty} \int_a^b f_n(x)\, d\alpha$,i.e., the series $\sum f_n$ is integrable term by term.

Proof. Let $\langle S_n \rangle$ denotes the sequence of partial sums of $\sum f_n$. Since $\sum f_n$ converges uniformly to f on [a, b], the sequence $\langle S_n \rangle$ converges uniformly to f. Then S_n being the sum of n integrable functions is integrable for each n. Therefore, by theorem 1, f is also integrable in Riemann sense and

$$\int_a^b f(x)\,dx = \lim_{n\to\infty} \int_a^b S_n(x)\,dx$$

But
$$\int_a^b S_n(x)dx = \int_a^b f_1(x)dx + \int_a^b f_2(x)\,dx + \cdots + \int_a^b f_n(x)dx$$

$$= \sum_{i=1}^{n} \int_a^b f_i(x)\,dx$$

$$\int_a^b f(x)\,dx = \lim_{n\to\infty} \sum_{i=1}^{n} \int_a^b f_i(x)\,dx$$

$$= \sum_{i=1}^{\infty} \int_a^b f_i(x)\,d\alpha$$

and the proof of the theorem is complete.

Example 1. Consider the sequence $\langle f_n \rangle$ for which $f_n(x) = nxe^{-nx^2}$, $n \in N$, $x \in [0,1]$. We note that

$$f(x) = \lim_{n\to\infty} f_n(x)$$

$$= \lim_{n\to\infty} \frac{nx}{1 + \frac{nx^2}{1!} + \frac{n^2x^4}{2!} + \ldots\ldots} = 0, \quad x \in (0,1]$$

Then

$$\int_0^1 f(x)\,dx = 0$$

$$\int_0^1 f_n(x)dx = \int_0^1 nxe^{-nx^2}dx$$

$$= \frac{1}{2}\int_0^n e^{-t}dt, \;\; t = nx^2$$

$$= \frac{1}{2}\left[1 - e^{-n}\right]$$

Therefore

$$\lim_{n\to\infty} \int f_n(x)dx = \lim_{n\to\infty} \frac{1}{2}\left[1 - e^{-n}\right] = \frac{1}{2}.$$

If $\langle f_n \rangle$ were uniformly convergent, then $\int_0^1 f(x)\,dx$ should have been equal to $\lim_{n\to\infty}\int f_n(x)\,dx$.

But it is not the case. Hence the given sequence is not uniformly convergent to f. In fact, x = 0 is the point of non-uniform convergence.

Example 2. Consider the series $\sum_{n=1}^{\infty}\frac{x}{(n+x^2)^2}$. This series is uniformly convergent and so is integrable term by term. Thus

$$\int_0^1\left(\sum_{n=1}^{\infty}\frac{x}{(n+x^2)^2}\right)dx = \lim_{m\to\infty}\sum_{n=1}^{m}\int_0^1\frac{x}{(n+x^2)^2}$$

$$= \lim_{m\to\infty}\sum_{n=1}^{m}\int_0^1 x(n+x^2)^{-2}dx$$

$$= \lim_{m\to\infty}\sum_{n=1}^{m}\left[\frac{(n+x^2)^{-1}}{-2}\right]_0^1$$

$$= \lim_{m\to\infty}\sum_{n=1}^{m}\frac{1}{2}\left(\frac{1}{n}-\frac{1}{n+1}\right)$$

$$= \lim_{m\to\infty}\frac{1}{2}\left[\left(1-\frac{1}{2}\right)+\left(\frac{1}{2}-\frac{1}{3}\right)+\ldots+\left(\frac{1}{m}-\frac{1}{m+1}\right)\right]$$

$$= \lim_{m\to\infty}\frac{1}{2}\left(1-\frac{1}{m+1}\right)=\frac{1}{2}$$

Example 3. Consider the series $\sum_{n=1}^{\infty}\left[\frac{nx}{(1+n^2x^2)}-\frac{(n-1)x}{(1+(n-1)^2x^2)}\right], \quad a\le x\le 1.$

Let $S_n(x)$ denote the partial sum of the series. Then

$$S_n(x)=\frac{nx}{(1+n^2x^2)}$$

and so $$f(x)=\lim_{n\to\infty}S_n(x)=0 \quad \text{for all } x\in[0,1]$$

As we know that 0 is point of non-uniform convergence of the sequence $\langle S_n(x)\rangle$, the given series is not uniformly convergent on [0, 1]. But

$$\int_0^1 f(x)\,dx = \int_0^1 0\,dx = 0$$

and
$$\int_0^1 S_n(x)\,dx = \int_0^1 \frac{nx}{(1+n^2x^2)}\,dx$$

$$= \frac{1}{2n}\int_0^1 \frac{2n^2x}{(1+n^2x^2)}\,dx$$

$$= \frac{1}{2n}\left[\log\left(1+n^2x^2\right)\right]_0^1$$

$$= \frac{1}{2n}\left[\log\left(1+n^2\right)\right].$$

Hence

$$\lim_{n\to\infty}\int_0^1 S_n(x)\,dx = \lim_{n\to\infty}\frac{1}{2n}\left[\log\left(1+n^2\right)\right]\left(\frac{\infty}{\infty}\text{ form}\right)$$

$$= \lim_{n\to\infty}\frac{n}{1+n^2}\left(\frac{\infty}{\infty}\text{ form}\right)$$

$$= \lim_{n\to\infty}\frac{1}{2n} = 0.$$

Thus

$$\int_0^1 f(x)\,dx = \lim_{n\to\infty}\int_0^1 S_n(x)\,dx,$$

and so the series is integrable term by term although 0 is a point of non-uniform convergence.

Theorem 3. Let $\{g_n\}$ be a sequence of functions of bounded variation on [a, b] such that $g_n(a) = 0$, and suppose that there is a function g such that

$$\lim_{n\to\infty} V(g - g_n) = 0$$

and g(a) = 0. Then for every continuous function f on [a, b], we have

$$\lim_{n\to\infty}\int_a^b f\,dg_n = \lim_{n\to\infty}\int_a^b f\,dg$$

and $g_n \to g$ uniformly on [a, b].

Proof. If V denotes the total variation on [a, b], then

$$V(g) \le V(g_n) + V(g - g_n)$$

Since g_n is of bounded variation and $\lim_{n\to\infty} V(g - g_n) = 0$ it follows that total variation of g is finite and so g is of bounded variation on [a, b]. Thus the integrals in the assertion of the theorem exist.

Suppose $|f(x) \le M|$ on [a, b]. Then

$$\left|\int_a^b f\ dg - \int_a^b f\ dg_n\right| = \left|\int_a^b f\ d(g - g_n)\right|$$

$$\le M\ V(g - g_n).$$

Since $V(g - g_n) \to 0$ as $n \to \infty$, it follows that

$$\int_a^b f\ dg = \lim_{n\to\infty} \int_a^b f\ dg_n.$$

Furthermore,

$$|g(x) - g_n(x)| \le V(g - g_n), \qquad a \le x \le b$$

Therefore, as $n \to \infty$, we have

$g_n \to g$ uniformly.

2.8. Uniform Convergence and Differentiation

If f and g are derivable, then

$$\frac{d}{dx}\left[f(x) + g(x)\right] = \frac{d}{dx} f(x) + \frac{d}{dx} g(x)$$

and that this can be extended to finite number of derivable functions.

In this section, we shall extend this phenomenon under some suitable condition to infinite number of functions.

Theorem 1. Suppose $\{f_n\}$ is a sequence of functions, differentiable on [a, b] and such that $\{f_n(x_0)\}$ converges for some point x_0 on [a, b]. If $\{f_n'\}$ converges uniformly on [a, b], then $\{f_n\}$ converges uniformly on [a, b], to a function f, and

$$f'(x) = \lim_{n\to\infty} f_n'(x) \ \ (a \le x \le b).$$

Proof. Let $\varepsilon > 0$ be given. Choose N such that $n \ge N,\ m \ge N$ implies

$$|f_n(x_0) - f_m(x_0)| < \frac{\varepsilon}{2} \qquad (1)$$

and

$$\left|f_n'(t)-f_m'(t)\right| < \frac{\varepsilon}{2(b-a)}\ (a \le t < b) \qquad (2).$$

Application of mean value theorem to the function $f_n - f_m$, (2) yields

$$\left|f_n(x)-f_m(x)-f_n(t)+f_m(t)\right| \le \frac{|x-t|\varepsilon}{2(b-a)} \le \frac{\varepsilon}{2} \qquad (3)$$

for any x and t on [a, b] if $n \ge N,\ m \ge N$. Since

$$\left|f_n(x)-f_m(x)\right| \le \left|f_n(x)-f_m(x)-f_n(x_0)+f_m(x_0)\right| + \left|f_n(x_0)-f_m(x_0)\right|.$$

the relation (1) and (3) imply for $n \ge N,\ m \ge N$,

$$\left|f_n(x)-f_m(x)\right| < \varepsilon/2+\varepsilon/2 = \varepsilon\ (a \le x < b).$$

Hence, by Cauchy criterion for uniform convergence, it follows that $\{f_n\}$ converges uniformly on [a, b]. Let

$$f(x) = \lim_{n\to\infty} f_n(x) \qquad (a \le x \le b).$$

For a fixed point $x \in [a,b]$, let us define

$$\phi_n(t) = \frac{f_n(t)-f_n(x)}{t-x}, \qquad \phi(t) = \frac{f(t)-f(x)}{t-x} \qquad (4)$$

for $a \le t \le b,\ t \ne x$. Then

$$\lim_{t\to x}\phi_n(t) = \lim_{t\to x}\frac{f_n(t)-f_n(x)}{t-x} = f_n'(x) \qquad (n = 1, 2,\ldots) \qquad (5)$$

Further, (3) implies

$$\left|\phi_n(t)-\phi_m(t)\right| \le \frac{\varepsilon}{2(b-a)}\ n \ge N,\ m \ge N.$$

Hence $\{\phi_n\}$ converges uniformly for $t \ne x$. We have proved just now that $\{f_n\}$ converges to f uniformly on [a, b]. Therefore (4) implies that

$$\lim_{n\to\infty}\phi_n(t) = \phi(t) \qquad (6)$$

uniformly for $(a \le t < b)$, $t \ne x$. Therefore using uniform convergence of $\langle\phi_n\rangle$ and (5), we have

$$\lim_{t\to x}\phi(t) = \lim_{t\to x}\lim_{n\to\infty}\phi_n(t)$$

$$= \lim_{n\to\infty}\lim_{t\to x}\phi_n(t)$$

$$= \lim_{n\to\infty} f_n'(x).$$

But $\lim_{t\to x} \phi(t) = f'(x)$. Hence

$$f'(x) = \lim_{n\to\infty} f_n'(x).$$

Remark 1. If in addition to the above hypothesis, each f_n' is continuous, then the proof becomes simpler. Infact, we have then

Theorem 2. Let $\langle f_n \rangle$ be a sequence of functions such that

(i) each f_n is differentiable on [a, b].

(ii) each f_n' is continuous on [a, b].

(iii) $\langle f_n \rangle$ converges to f on [a, b].

(iv) $\langle f_n' \rangle$ converges uniformly to g on [a, b], then f is differentiable and $f_n'(x) = g(x)$ for all $x \in [a,b]$.

Proof. Since each f_n' is continuous on [a, b] and $\langle f_n' \rangle$ converges uniformly to g on [a, b], the application of Theorem 1 of section 2.6 of this unit implies that g is continuous and hence Riemann integrable. Therefore, Theorem 1 of section 2.7 of this unit implies

$$\int_a^t g(x)\,dx = \lim_{n\to\infty} \int_a^t f_n'(x)\,dx$$

But, by Fundamental theorem of integral calculus,

$$\int_a^t f_n'(x)\,dx = f_n(t) - f_n(a)$$

Hence

$$\int_a^t g(x)\,dx = \lim_{n\to\infty} \left[f_n(t) - f_n(a) \right]$$

Since $\langle f_n \rangle$ converges to f on [a, b], we have

$$\lim_{n\to\infty} f_n(t) = f(t) \text{ and } \lim_{n\to\infty} f_n(a) = f(a).$$

Hence

$$\int_a^t g(x)\,dx = f(t) - f(a)$$

and so

$$\frac{d}{dt}\left(\int_a^t g(x)\,dx\right) = f'(t)$$

or $$g(t) = f'(t),\ t \in [a,b].$$

This completes the proof of the theorem.

The series version of Theorem 2 is

Theorem 3. If a series $\sum f_n$ converges to f on [a, b] and

(i) each f_n is differentiable on [a, b]

(ii) each f_n' is continuous on [a, b]

(iii)the series $\sum f_n'$ converges uniformly to g on [a, b]

then f is differentiable on [a, b] and $f'(x) = g(x)$ for all $x \in [a,b]$.

Proof. Let $\langle S_n \rangle$ be the sequence of partial sums of the series $\sum_{n=1}^{\infty} f_n$. Since $\sum f_n$ converges to f on [a, b], the sequence $\langle S_n \rangle$ converges to f on [a, b]. Further, since $\sum f_n'$ converges uniformly to g on [a, b], the sequence $\langle S_n' \rangle$ of partial sums converges uniformly to g on [a, b].

Hence, theorem 2 is applicable and we have

$$f'(x) = g(x) \text{ for all } x \in [a,b].$$

Example 1. Consider the series $\sum_{n=1}^{\infty}\left[\frac{nx}{(1+n^2x^2)} - \frac{(n-1)x}{(1+(n-1)^2x^2)}\right]$.

For this series, we have

$$S_n(x) = \frac{nx}{(1+n^2x^2)}, \qquad 0 \le x \le 1$$

We have seen that 0 is a point of non-uniform convergence for this sequence. We have

$$f(x) = \lim_{n\to\infty} S_n(x) = \lim_{n\to\infty}\frac{nx}{(1+n^2x^2)}$$

$$= 0 \quad \text{for } 0 \le x \le 1.$$

Therefore

$$f'(0) = 0$$

$$S_n'(0) = \lim_{h\to\infty} \frac{S_n(0+h) - S_n(0)}{h}$$

$$= \lim_{h\to 0} \frac{n}{(1+n^2h^2)} = n$$

Hence

$$\lim_{n\to\infty} S_n'(0) = \infty.$$

Then

$$f'(0) \neq \lim_{n\to\infty} S_n'(0).$$

Example 2. Consider the series $\sum_{n=1}^{\infty} \frac{\sin nx}{n^3}$, $x \in R$. We have

$$f_n(x) = \frac{\sin nx}{n^3}$$

$$f_n'(x) = \frac{\cos nx}{n^2}.$$

Thus

$$\sum f_n'(x) = \sum \frac{\cos nx}{n^2}$$

Since $\left|\frac{\cos nx}{n^2}\right| \leq \frac{1}{n^2}$ and $\sum \frac{1}{n^2}$ is convergent, therefore, by Weierstrass's M-test the series $\sum f_n'(x)$ is uniformly as well as absolutely convergent for all $x \in R$ and so $\sum f_n$ can be differentiated term by term.

Hence
$$\left(\sum_{n=1}^{\infty} f_n\right)' = \sum_{n=1}^{\infty} f_n'$$

or
$$\left(\sum_{n=1}^{\infty} \frac{\sin nx}{n^3}\right)' = \sum_{n=1}^{\infty} \frac{\cos nx}{n^2}$$

2.9 Weierstrass's Approximation Theorem

Weierstrass proved an important result regarding approximation of continuous function which has many applications in numerical methods and other branches of mathematics.

The following computation shall be required for the proof of Weierstrass's approximation theroem.

For any $p, q \in R$, we have, by Binomial Theorem

$$\sum_{k=0}^{n}\binom{n}{k}p^k q^{n-k} = (p+q)^n,, \qquad n \in I, \qquad (1)$$

where

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$

Differentiating with respect to p, we obtain

$$\sum_{k=0}^{n}\binom{n}{k}kp^{k-1}q^{n-k} = n(p+q)^{n-1},$$

which implies

$$\sum_{k=0}^{n}\frac{k}{n}\binom{n}{k}p^k q^{n-k} = p(p+q)^{n-1}, \quad n \in I \qquad (2).$$

Differentiating once more, we have

$$\sum_{k=0}^{n}\frac{k^2}{n}\binom{n}{k}p^{k-1}q^{n-k} = p(n-1)(p+q)^{n-2} + (p+q)^{n-1}$$

and so

$$\sum_{k=0}^{n}\frac{k^2}{n^2}\binom{n}{k}p^k q^{n-k} = p^2(1-\frac{1}{n})(p+q)^{n-2} + \frac{p}{n}(p+q)^{n-1} \qquad (3).$$

Now if $x \in [0,1]$, take p = x and q = 1-x. Then (1), (2) and (3) yield

$$\begin{cases} \sum_{k=0}^{n}\binom{n}{k}x^k(1-x)^{n-k} = 1 \\ \sum_{k=0}^{n}\frac{k}{n}\binom{n}{k}x^k(1-x)^{n-k} = x \\ \sum_{k=0}^{n}\frac{k^2}{n^2}\binom{n}{k}x^k(1-x)^{n-k} = x^2\left(1-\frac{1}{n}\right)+\frac{x}{n} \end{cases} \qquad (4).$$

On expanding $\left(\frac{k}{n}-x\right)^2$, it follows from (4) that

$$\sum_{k=0}^{n}\left(\frac{k}{n}-x\right)^2\binom{n}{k}x^k(1-x)^{n-k} = \frac{x(1-x)}{n} \qquad 0 \le x \le 1 \qquad (5).$$

For any $f \in [0,1]$, we define a sequence of polynomials $\{B_n\}_{n=1}^{\infty}$ as follows:

$$B_n(x)=\sum_{k=0}^{n}\binom{n}{k}x^k(1-x)^{n-k} f\left(\frac{k}{n}\right), \qquad 0\le x\le 1,\ n\in I \qquad (6).$$

The polynomial B_n is called the nth **Bernstein Polynomial** for f.

We are in a position to state and prove Weierstrass's Theorem.

Theorem 1 (Weierstrass's Approximation Theorem). If f is real continuous function defined on [a,b] then there exists a sequence of real polynomials $\{P_n\}$ which converges uniformly to f(x) on [a,b] i.e., $\lim_{n\to\infty} P_n(x)=f(x)$ uniformly on [a,b].

Proof. If a = b, then f(x) = f(a).

Then, the theorem is true by taking $P_n(x)$ to be a constant polynomial defined by

$$P_n(x)=f(a)\ \forall n$$

Thus we assume that a < b

$f=\dfrac{x-a}{b-a}$ is continuous mapping of [a,b] onto [0,1].

So, in our discussion W.L.O.G. we take a = 0, b = 1.

Now we know that for positive integer n and k where $0\le k\le n$, the binomial coefficients $\binom{n}{k}$ i.e,

n_{c_k} is defined as $\dbinom{n}{k}=\dfrac{n!}{k!(n-k)!}$

Now, we define the polynomial B_n where

$$B_n(x)=\sum\binom{n}{k}x^k(1-x)^{n-k} f\left(\frac{k}{n}\right) \qquad (*)$$

The polynomial defined in (*) is called Bernstain polynomial as shown in above equation (6).

We shall prove that certain Bernstain polynomial exists which uniformly converges to f on [0,1].

Now consider the identity

$$\sum_{k=0}^{n}\binom{n}{k}x^k(1-x)^{n-k}=\left[x+(1-x)\right]^n=1 \qquad (1)$$

[This is the binomial exp. of $x+[1-x]^n$]

Differentiating w.r.t. x, we get

$$\sum_{k=0}^{n}\binom{n}{k}\left[kx^{k-1}(1-x)^{n-k}-(n-k)x^{k}(1-x)^{n-k-1}\right]=0$$

$$\Rightarrow\sum_{k=0}^{n}\binom{n}{k}x^{k-1}(1-x)^{n-k-1}(k-nx)=0$$

Multiplying by x(1-x) yields

$$\Rightarrow\sum_{k=0}^{n}\binom{n}{k}\left[x^{k}(1-x)^{n-k}(k-nx)\right]=0 \qquad (2)$$

Differentiating again w.r.t. x, we get

$$\sum_{k=0}^{n}\binom{n}{k}\left[-nx^{k}(1-x)^{n-k}+x^{k-1}(1-x)^{n-k-1}(k-nx)^{2}\right]=0$$

which on applying (1),we get

$$\sum_{k=0}^{n}\binom{n}{k}\left[x^{k-1}(1-x)^{n-k-1}(k-nx)^{2}\right]=n$$

Multiplying by x(1-x), we get

$$\sum_{k=0}^{n}\binom{n}{k}x^{k}(1-x)^{n-k}(k-nx)^{2}=nx(1-x)$$

$$\Rightarrow\sum_{k=0}^{n}\binom{n}{k}x^{k}(1-x)^{n-k}\left(x-\frac{k}{n}\right)^{2}=\frac{x(1-x)}{n} \qquad (3)$$

Since the maximum value of x(1-x) in [0,1] is ¼.

$$f(x)=x(1-x),\ f'(x)=1-2x$$
$$\Rightarrow f'(x)=0\Rightarrow 1-2x=0$$
$$\Rightarrow x=1/2\Rightarrow f(1/2)=1/4.$$

So, (3) can be written as

$$\Rightarrow\sum_{k=0}^{n}\binom{n}{k}x^{k}(1-x)^{n-k}\left(x-\frac{k}{n}\right)^{2}\le\frac{1}{4n} \qquad (4)$$

Now f is continuous on [0,1]. So, f is bounded and uniformly continuous on [0,1].

$\Rightarrow\exists\ K>0$ such that

$$|f(x)|\le K\quad\forall x\in[0,1]$$

and by uniform continuity for given $\varepsilon>0$, there exists $\delta>0$ such that for all $x\in[0,1]$.

$$\Rightarrow \left| f(x) - f\left(\frac{k}{n}\right) \right| < \frac{\varepsilon}{2} \text{ whenever } \left| x - \frac{k}{n} \right| < \delta \qquad (5).$$

Now for any fixed but arbitrary x in [0,1], then n- values 0,1,2,………..,n of k can be divided into two parts as follows:

Let A be the set of values of k for which $\left| x - \frac{k}{n} \right| < \delta$ and B be the set of remaining values for which $\left| x - \frac{k}{n} \right| \geq \delta$.

Now for $k \in B$, we get by (4)

$$\sum_{k \in B} \binom{n}{k} x^k (1-x)^{n-k} \delta^2 \leq \sum_{k \in B} \binom{n}{k} x^k (1-x)^{n-k} \left(x - \frac{k}{n} \right)^2 \leq \frac{1}{4n}$$

$$\Rightarrow \sum_{k \in B} \binom{n}{k} x^k (1-x)^{n-k} \leq \frac{1}{4n\delta^2} \qquad (6)$$

Now

$$|f(x) - B_n(x)| = \left| 1.f(x) - \sum_{k=0}^{n} \binom{n}{k} x^k (1-x)^{n-k} f\left(\frac{k}{n}\right) \right|$$

(By (1))

$$= \left| \sum_{k=0}^{n} \binom{n}{k} x^k (1-x)^{n-k} \left[f(x) - f\left(\frac{k}{n}\right) \right] \right|$$

$$|f(x) - B_n(x)| \leq \sum_{k=0}^{n} \binom{n}{k} x^k (1-x)^{n-k} \left| f(x) - f\left(\frac{k}{n}\right) \right|$$

We split the summation on R.H.S into two parts accordingly as

$$\left| x - \frac{k}{n} \right| < \delta \text{ or } \left| x - \frac{k}{n} \right| \geq \delta$$

Let $k \in A$ or $k \in B$.

Thus we have

$$|f(x) - B_n(x)| \leq \sum_{k \in A} \binom{n}{k} x^k (1-x)^{n-k} \left| f(x) - f\left(\frac{k}{n}\right) \right| + \sum_{k \in B} \binom{n}{k} x^k (1-x)^k \left| f(x) - f\left(\frac{k}{n}\right) \right|$$

$$< \frac{\varepsilon}{2} \sum_{k \in A} \binom{n}{k} x^k (1-x)^{n-k} + 2K \sum_{k \in B} \binom{n}{k} x^k (1-x)^{n-k}$$

$$< \frac{\varepsilon}{2} + \frac{2K}{4n\delta^2} < \varepsilon \text{ for all values of } n > \frac{K}{\varepsilon\delta^2}.$$

Thus $\{B_n(x)\}$ converges uniformly to f(x) on [0, 1].

Hence the proof.

Example1. If f is continuous on [0,1] and if $\int x^n f(x)dx = 0$ for n=0,1,2,………. Then show that f(x) = 0 on [0,1].

Solution. Let $p(x) = a_0 + a_1 x + a_2 x^2 + \dots\dots\dots\dots + a_n x^n$ be a polynomial with real co-efficients defined on [0,1], then

$$\int_0^1 p(x)f(x)dx = \int_0^1 \left(\sum_{n=0}^{\infty} a_n x^n\right) f(x)dx$$

$$= \sum_{n=0}^{\infty} a_n \int_0^1 x^n f(x)dx = \sum_{n=0}^{\infty} a_n .0 = 0.$$

Thus the integral of product of f with any polynomial is zero.

Now, since f is continuous on [0,1], therefore by Weierstrass's approximation theorem, there exists a seq. $\{p_n\}$ of real polynomial such that $p_n \to f$ uniformly on [0,1].

$\Rightarrow p_n f \to f^2$ is uniformly on [0,1]

Since f being continuous and bounded on [0,1], therefore

$$\int_0^1 f^2 dx = \lim_{n\to\infty} \int_0^1 p_{n.} .f dx = 0$$

Therefore, $f^2(x) = 0$ on [0,1].

Hence f(x) = 0 on [0,1].

2.10 References

1. Walter Rudin, Principles of Mathematical Analysis (3rd edition) McGraw-Hill, Kogakusha, 1976, International Student Edition.
2. T. M. Apostol, Mathematical Analysis, Narosa Publishing House, New Delhi, 1974.
3. H.L. Royden, Real Analysis, Macmillan Pub. Co., Inc. 4th Edition, New York, 1993.
4. G. De Barra, Measure Theory and Integration, Wiley Eastern Limited, 1981.
5. R.R. Goldberg, Methods of Real Analysis, Oxford & IBH Pub. Co. Pvt. Ltd, 1976.
6. R. G. Bartle, The Elements of Real Analysis, Wiley International Edition, 2011.
7. S.C. Malik and Savita Arora, Mathematical Analysis, New Age International Limited, New Delhi, 2012.

UNIT – III

POWER SERIES AND FUNCTION OF SEVERAL VARIABLES

Structure

3.0 Introduction

In this unit, we study convergence and divergence of a power series and applications of Abel's theorem. Tauber showed that the converse of Abel's theorem can be obtained by imposing additional condition on coefficients, whenever the converse of Abel's theorem is false in general. Many of the concepts i.e., continuity, differentiability, chain rule, partial derivatives etc are extended to functions of more than one independent variable.

3.1 Unit Objectives

After going through this unit, one will be able to

- understand the concept of power series and radius of convergence.
- identify the notation associated with functions of several variables
- familiar with the chain rule, partial derivatives and concept of derivation in an open subset of R^n.
- know the features of Young and Schwarz's Theorems.

3.2 Power Series

A very important class of series to study is power series. A power series is a type of series with terms involving a variable. Evidently, if the variable is x, then all the terms of the series involve powers of x. So we can say that a power series can be design of as an infinite polynomial. In this section we will give the definition of the power series as well as the definition of the radius of convergence, uniform convergence and uniqueness theorem, Abel and Tauber theorems.

Definition 1. A power series is an infinite series of the form $\sum_{n=0}^{\infty} a_n x^n$ where $a_n's$ are called its coefficients.

Definition 2 (Convergence of power series). It is clear that for x = 0, every power series is convergent, independent of the values of the coefficients. Now, we are given three possible cases about the convergence of a power series.

(a) The series converges for only x = 0 which is trivial point of convergence, then it is called "nowhere convergent"

e.g. $\sum n! x^n$ converges only for x = 0 and for $x \neq 0$, we have

$$\lim_{n\to\infty} n! x^n = \infty.$$

Thus the terms of the series do not converge for $x \neq 0$ and thus the series converges only for x = 0. Hence it is 'Nowhere convergent' series.

(b) The series converges absolutely for all values of x, then it is called "Everywhere convergent".

e.g. The series converges absolutely for all values of x,

$$u_n = \frac{x^n}{n!}, u_{n+1} = \frac{x^{n+1}}{n+1!}$$

$$\lim_{n\to\infty}\left|\frac{u_n}{u_{n+1}}\right| = \left|\frac{x^n}{n!} \times \frac{(n+1)!}{x^{n+1}}\right| = \left|\frac{n+1}{x}\right| = \infty.$$

By D-Ratio test, the series converges for all values of x. So, it is called "Everywhere convergent" series.

(c) The series converges for some values of x and diverges for others.

e.g. The series $\sum_{n=0}^{\infty} x^n$ converges for x < 1 and diverges for x > 1.

The collection of points x for which the series is convergent is called its "Region of convergence".

Definition 3. Let $\sum_{n=0}^{\infty} a_n x^n$ be a power series. Then, applying Cauchy's root test, we observe that the power series $\sum_{n=0}^{\infty} a_n x^n$ is convergent if

$$|x| < \frac{1}{L},$$

where

$$L = \overline{\lim} |a_n|^{1/n} .$$

The series is divergent if $|x| > \frac{1}{L}$.

Taking

$$R = \frac{1}{\overline{\lim} |a_n|^{1/n}} .$$

We will prove that the power series is absolutely convergent if $|x| < R$ and divergent if $|x| > R$. If $a_0, a_1, \ldots.$ are all real and if x is real, we get an interval $-R < x < R$ inside which the series is convergent.

If x is replaced by a complex number z, the power series $\sum_{n=0}^{\infty} a_n z^n$ converges absolutely at all points z inside the circle $|z| = R$ and does not converge at any point outside this circle. The circle is known as **circle of convergence** and **R** is called **radius of convergence**. In case of real power series, the interval (-R, R) is called interval of convergence.

If $\overline{\lim} |a_n|^{1/n} = 0,$ then $R = \infty$ and the power series converges for all finite values of x. The function represented by the sum of series is then called an **Entire function** or an integral function. For example, $e^z, \sin z$ and $\cos z$ are integral functions.

If $\overline{\lim} |a_n|^{1/n} = \infty,$ $R = 0,$ the power series does not converge for any value of x except x = 0.

Definition 4. Let R be the radius of convergence of the power series $\sum_{n=0}^{\infty} a_n z^n$, then the open interval (-R, R) is called **the interval of convergence** for the given power series.

Theorem 1. Let $\sum a_n x^n$ be a power series such that $\overline{\lim_{n\to\infty}} |a_n|^{1/n} = \frac{1}{R}$.

Then the power series is convergent with radius of convergence R.

Proof. Given that

$$\overline{\lim_{n\to\infty}}|a_n|^{1/n} = \frac{1}{R}$$

So, $$\overline{\lim_{n\to\infty}}|a_n x^n|^{1/n} = \frac{|x|}{R}$$

Hence by Cauchy's Root test, the series $\sum a_n x^n$ is convergence if $\frac{|x|}{R} < 1$ and divergent if $\frac{|x|}{R} > 1$ i.e, convergent if $|x| < R$ and divergent if $|x| > R$. Hence by definition, R is radius of convergence of the given power series.

Remark 1. (i) From the proof of above theorem, it follows that if for the series $\sum a_n x^n$,

$$\overline{\lim_{n\to\infty}}|a_n|^{1/n} = \frac{1}{R}$$

then the series is absolutely convergent.

(ii) In view of the last theorem, we define the power series of convergence in the following way:

Consider the power series $\sum a_n x^n$, then the radius of convergence of this series is given by

$$R = \frac{1}{\overline{\lim}|a_n|^{1/n}} \text{ when } \overline{\lim}|a_n|^{1/n} > 0$$

$$= 0 \text{ when } \overline{\lim}|a_n|^{1/n} = \infty$$

$$= \infty \text{ when } \overline{\lim}|a_n|^{1/n} = 0.$$

Obviously $R = \infty$ for an "**everywhere convergent**" and $R = 0$ for a "**nowhere convergent**" series.

Theorem 2. If a power series $\sum a_n x^n$ converges for $x = x_0$ then it is absolutely convergent for every $x = x_1$ where $|x_1| < |x_0|$.

Proof. Given that the series $\sum a_n x^n$ is convergent.

Thus $a_n x_0^n \to 0$ as $n \to \infty$.

Hence for $\varepsilon = 1/2$ (say), there exists an integer N such that

$$|a_n x_0^n| < \frac{1}{2} \forall n \geq N$$

Thus, we have

$$|a_n x_1^n| = |a_n x_0^n| \left|\frac{x_1}{x_0}\right|^n$$

$$< \frac{1}{2}\left|\frac{x_1}{x_0}\right|^n \ \forall n \geq N \qquad (*)$$

Now $|x_1| < |x_0| \Rightarrow \left|\frac{x_1}{x_0}\right| < 1$.

Thus $\sum \left|\frac{x_1}{x_0}\right|^n$ is geometric series with common ratio less than 1. So, it is convergent. By comparison test, the series $\sum |a_n x_1^n|$ converges.

$\Rightarrow \sum a_n x^n$ is absolutely convergent for every $x = x_1$ where $|x_1| < |x_0|$.

Theorem 3. If a power series $\sum a_n x^n$ diverges for $x = x'$ then it diverges for every $x = x''$, where $|x''| > |x'|$.

Proof. Given that the series $\sum a_n x^n$ diverges at $x = x'$.

Let x'' be such that $|x''| > |x'|$.

Let if possible, the series is convergent for $x = x''$, then by theorem 2, it must be convergent for all x such that $|x| < |x''|$.

In particular, it must be convergent at x' which is contradiction to the given hypothesis.

Hence the series diverges for every $x = x''$, where $|x''| > |x'|$.

Definition 5 (Radius of Convergence).

For the power series $\sum a_n x^n$, the radius of convergence is also defined by the relation $R = \lim_{n\to\infty}\left|\frac{a_n}{a_{n+1}}\right|$, provided the limit exists.

This definition is commonly used for numerical purpose as illustrated below:

Find the radius of convergence of following:

(1) $x + \frac{x^2}{2!} + \frac{x^3}{3!} + \ldots\ldots\ldots\ldots\ldots$

(2) $1 + x + 2!x^2 + 3!x^3 + \ldots\ldots\ldots\ldots\ldots\ldots\ldots$

(3) $\frac{1}{2}x+\frac{1.3}{2.5}x^2+\frac{1.3.5}{2.5.8}x^3+................$

(4) $x+\frac{1}{2^2}x^2+\frac{1.2}{3^3}x^3+\frac{1.2.3}{4^4}x^4+..............$

Solution. (1) Here $a_n=\frac{1}{n!}$

$$R=\lim_{n\to\infty}\left|\frac{a_n}{a_{n+1}}\right|=\lim_{n\to\infty}\left|\frac{1}{n!}\times(n+1)!\right|=\infty$$

The series converges for all values of x i.e, everywhere convergent.

(2) Here $a_n=n!$

$$R=\lim_{n\to\infty}\left|\frac{n!}{(n+1)!}\right|=0$$

So, the series converges for no value of x other than zero. So, it is nowhere convergent series.

(3) Here $a_n=\frac{1.3.5..................(2n-1)}{2.5.8................(3n-1)}$

$$R=\lim_{n\to\infty}\left|\frac{1.3.5...............(2n-1)}{2.5.8...............(3n-1)}\times\frac{2.5.8.............(3n-1)(3n+2)}{1.3.5..............(2n-1)(2n+1)}\right|$$

$$=\lim_{n\to\infty}\left|\frac{3+2/n}{2+1/n}\right|=\frac{3}{2}$$

So series converges for all x where $|x|<\frac{3}{2}$.

(4) Here $a_n=\frac{(n-1)!}{n^n}$

$$R=\lim_{n\to\infty}\left|\frac{(n-1)!}{n^n}\times\frac{(n+1)^{n+1}}{n!}\right|=\lim_{n\to\infty}\left(1+\frac{1}{n}\right)^{n+1}=e.$$

So the series is convergent for all x where $|x|<e$.

Definition 6. Let f(x) be a function which can be express in terms of the power series as

$$f(x)=\sum_{n=0}^{\infty}a_nx^n,$$

then f(x) is called sum function of the power series $\sum a_nx^n$.

Remark 2. We have defined the uniform convergence of a series in a closed interval always. Thus, if a power series converges uniformly for $|x| < R$, then we must express this fact by saying that the series converges uniformly in closed interval $[-R+\varepsilon, R-\varepsilon]$, where $\varepsilon > 0$ may be arbitrary chosen, however if a power series converges absolutely for $|x| < R$, then we can directly say that the series converges absolutely in (-R, R).

Theorem 4. Suppose the series $\sum_{n=0}^{\infty} a_n x^n$ converges for $|x| < R$ and define

$$f(x) = \sum_{n=0}^{\infty} a_n x^n \qquad (|x| < R).$$

Then

(i) $\sum_{n=0}^{\infty} a_n x^n$ converges uniformly on $[-R+\in, R-\in], \in > 0.$

(ii) The function f in continuous and differentiable in (-R, R)

(iii) $f'(x) = \sum_{n=1}^{\infty} na_n x^{n-1} \qquad (|x| < R)$

Proof. (i) Let $\in$ be a positive number. If $|x| \le R - \in$, we have

$$|a_n x^n| \le |a_n (R-\in)^n|$$

Since every power series converges absolutely in interior of its interval of convergence by Cauchy's root test, the series $\sum a_n (R-\in)^n$ converges absolutely and so, by Weierstrass's M-test, $\sum a_n x^n$ converges uniformly on $[-R+\in, R-\in]$.

(ii) Also then the sum f(x) of $\sum a_n x^n$ is continuous and differentiable on (-R, R) and $\sum a_n x^n$ is uniformly convergent on $[-R+\varepsilon, R-\varepsilon]$.

Therefore, its sum function is continuous and differentiable on (-R, R).

(iii) Now consider the series $\sum na_n x^{n-1}$.

Since $(n)^{1/n} \to 1$ as $n \to \infty$, we have

$$\overline{\lim}(n\,|\,a_n\,|)^{1/n} = \overline{\lim}(|\,a_n\,|)^{1/n}$$

Hence the series $\sum a_n x^n$ and $\sum na_n x^{n-1}$ have the same interval of convergence. Since $\sum na_n x^{n-1}$ is a power series, it converges uniformly in $[-R+\varepsilon, R-\varepsilon]$ for every $\in > 0$. Then, by term by term differentiation yields

$$\sum na_n x^{n-1} = f'(x) \text{ if } |x| < R - \in.$$

But, given any x such that $|x| < R$ we can find an $\in > 0$ such that $|x| < R - \in$. Hence

$$\sum na_n x^{n-1} = f'(x) \text{ if } |x| < R.$$

Note. It follows from the above theorem 4 that by repeated application of the theorem f can be differentiable any number of time and series obtained by differentiation at each step has the same radius of convergence as series $\sum a_n x^n$.

Theorem 5. Under the hypothesis of Theorem 4, f has derivative of all orders in (-R, R) which are given by

$$f^{(k)}(x) = \sum_{n=k}^{\infty} n(n-1)(n-2)...(n-k+1)a_n x^{n-k} .$$

In particular

$$f^{(k)}(0) = \underline{|k}\, a_k, k = 0,1,2,..........$$

Proof. Let

$$f(x) = \sum_{n=0}^{\infty} na_n x^n .$$

Then by theorem 4,

$$f^{'}(x) = \sum_{n=1}^{\infty} n\, a_n x^{n-1}.$$

Again applying theorem 4 to $f^{'}(x)$, we have

$$f^{''}(x) = \sum_{n=2}^{\infty} n(n-1)a_n x^{n-2}$$

....

....

....

$$f^{(k)}(x) = \sum_{n=k}^{\infty} n(n-1)(n-2)......(n-k+1)\, a_n x^{n-k}.$$

Clearly $f^{(k)}(0) = \underline{|k}\, a_k$ the other sum vanish at x = 0.

Remark 3. If the coefficients of a power series are known, the values of the derivatives of f at the centre of the interval of convergence can be found from the relation

$$f^{(k)}(0) = \underline{|k}\, a_k .$$

Also we can find coefficient from the values at origin of $f, f', f'',...$

Theorem 6 (Uniqueness theorem). If $\sum a_n x^n$ and $\sum b_n x^n$ converge on some interval (-R, R), R > 0 to some function f, then

$$a_n = b_n \text{ for all } n \in \mathbf{N}.$$

Proof. Under the given condition, the function f have derivatives of all order in (-R, R) given by

$$f^{(k)}(x) = \sum_{n=k}^{\infty} n(n-1)(n-2)......(n-k+1)\, a_n x^{n-k}.$$

Putting x = 0, this yields

$$f^{(k)}(0) = k!a_k \text{ and } f^{(k)}(0) = k!b_k.$$

for all $k \in \mathbf{N}$. Hence

$$a_k = b_k \text{ for all } k \in \mathbf{N}.$$

This completes the proof of the theorem.

Theorem 7 (Abel's Theorem (First form)). If a power series $\sum_{n=o}^{\infty} a_n x^n$ converges at the point R of the interval of convergence (-R, R), then it uniformly converges in the interval [0, R].

Proof. Consider the sum

$$S_{n,p} = a_{n+1}R^{n+1} + a_{n+2}R^{n+2} + ...+ a_{n+p}R^{n+p}; p = 1, 2, ...$$

Then, we have

$$S_{n,1} = a_{n+1}R^{n+1}$$
$$S_{n,2} = a_{n+1}R^{n+1} + a_{n+2}R^{n+2}$$

and so on.

This gives

$$\left.\begin{aligned} a_{n+1}R^{n+1} &= S_{n,1} \\ a_{n+2}R^{n+2} &= S_{n,2} - S_{n,1} \\ &.............................. \\ &............................ \\ a_{n+p}R^{n+p} &= S_{n,p} - S_{n,p-1} \end{aligned}\right\} \quad (1)$$

Let ϵ > 0 be given.

Now the series $\sum_{n=0}^{\infty} a_n x^n$ is convergent at x = R.

The series of numbers $\sum_{n=0}^{\infty} a_n R^n$ is convergent and hence by Cauchy's general principle of convergence, there exists an integer N such that

$$\left|a_{n+1}R^{n+1} + a_{n+2}R^{n+2} + \ldots\ldots + a_{n+q}R^{n+q}\right| < \varepsilon \;\forall\; n \geq N \;\; \forall\, q = 1, 2, \ldots\ldots\ldots\ldots$$

$$\Rightarrow \quad \left|S_{n,q}\right| < \varepsilon \;\forall\; n \geq N \;\&\; q = 1, 2, \ldots\ldots\ldots\ldots \qquad (2)$$

Now if we take $x \in [0, R]$ i.e, $0 \leq x \leq R$, then we have

$$\left(\frac{x}{R}\right)^{n+p} \leq \left(\frac{x}{R}\right)^{n+p-1} \leq \ldots\ldots\ldots\ldots \leq \left(\frac{x}{R}\right)^{n+1} \leq 1. \qquad (3)$$

Now, consider for all $n \geq N$,

$$\left|a_{n+1}x^{n+1} + a_{n+2}x^{n+2} + \ldots\ldots\ldots\ldots\ldots + a_{n+p}x^{n+p}\right|$$

$$= \left|a_{n+1}R^{n+1}\left(\frac{x}{R}\right)^{n+1} + a_{n+2}R^{n+2}\left(\frac{x}{R}\right)^{n+2} + \ldots\ldots\ldots\ldots + a_{n+p}R^{n+p}\left(\frac{x}{R}\right)^{n+p}\right|$$

$$= \left|S_{n,1}\left(\frac{x}{R}\right)^{n+1} + \left(S_{n,2} - S_{n,1}\right)\left(\frac{x}{R}\right)^{n+2} + \ldots\ldots\ldots\ldots + \left(S_{n,p} - S_{n,p-1}\right)\left(\frac{x}{R}\right)^{n+p}\right|$$

$$= \left|S_{n,1}\left\{\left(\frac{x}{R}\right)^{n+1} - \left(\frac{x}{R}\right)^{n+2}\right\} + S_{n,2}\left\{\left(\frac{x}{R}\right)^{n+2} - \left(\frac{x}{R}\right)^{n+3}\right\} + \ldots\ldots\ldots + S_{n,p}\left(\frac{x}{R}\right)^{n+p}\right|$$

$$\leq \left|S_{n,1}\right|\left\{\left(\frac{x}{R}\right)^{n+1} - \left(\frac{x}{R}\right)^{n+2}\right\} + \left|S_{n,2}\right|\left\{\left(\frac{x}{R}\right)^{n+2} - \left(\frac{x}{R}\right)^{n+3}\right\} + \ldots\ldots\ldots + \left|S_{n,p}\right|\left(\frac{x}{R}\right)^{n+p}$$

$$< \varepsilon\left\{\left(\frac{x}{R}\right)^{n+1} - \left(\frac{x}{R}\right)^{n+2} + \left(\frac{x}{R}\right)^{n+2} - \left(\frac{x}{R}\right)^{n+3} + \ldots\ldots\ldots\ldots + \left(\frac{x}{R}\right)^{n+p-1} - \left(\frac{x}{R}\right)^{n+p} + \left(\frac{x}{R}\right)^{n+p}\right\}$$

$$< \varepsilon\left(\frac{x}{R}\right) < \varepsilon.1 = \varepsilon \qquad \text{(by (3))}$$

Thus we have proved that

$$\left|a_{n+1}x^{n+1} + a_{n+2}x^{n+2} + \ldots\ldots\ldots + a_{n+p}x^{n+p}\right| < \varepsilon \;\forall\; p \geq 1, \;\forall x \in [0, R].$$

Hence by Cauchy's criterion of convergence of series, the series $\sum_{n=0}^{\infty} a_n x^n$ converges uniformly on [0, R].

Remark 4. (i) In case, a power series with interval of convergence (-R, R) converges at $x = -R$, then the series is uniformly convergent in [-R, 0].

Similarly, if a series convergent at the end points -R and R, then the series is uniformly convergent on [-R, R].

(ii) If a power series with interval of convergence (-R, R) diverges at end point $x = R$, then it cannot be uniformly convergent on [0, R].

For, if the series is uniformly convergent on [0, R], it will converge at x = R. A contradiction to the given hypothesis.

Theorem 8 (Abel's theorem (second form)). Let $\sum_{n=0}^{\infty} a_n x^n$ be a power series with finite radius of convergence R and let $f(x) = \sum a_n x^n$; $|x| < R$. If the series $\sum a_n x^n$ converges at end point x = R then $\lim_{x \to R^-} f(x) = \sum a_n R^n$.

Proof. First we show that there is no loss of generality if we take $R = 1$.

$$\sum a_n x^n = \sum a_n R^n y^n = \sum b_n y^n \quad \text{where } b_n = a_n R^n.$$

Now, this is a power series with radius R', where

$$R' = \frac{1}{\overline{\lim}\left|a_n R^n\right|^{1/n}} = \frac{1}{\overline{\lim}\left|a_n\right|^{1/n} R} = \frac{R}{R} = 1.$$

So, if any series is given, we can transform it in another power series with unit radius of convergence. Hence we can take R = 1.

Thus, now it is sufficient to prove that let $\sum_{n=0}^{\infty} a_n x^n$ be a power series with unit radius of convergence and let $f(x) = \sum a_n x^n; |x| < 1$, if the series $\sum a_n$ converges then $\lim_{x \to 1^-} f(x) = \sum_{n=0}^{\infty} a_n$. .Let us proceed to prove the same.

$$S_n = a_0 + a_1 + \dots\dots\dots\dots + a_n$$

$$S_{-1} = 0 \text{ and } \sum_{n=0}^{\infty} a_n = S.$$

Then $$\sum_{n=0}^{m} a_n x^n = \sum_{n=0}^{m} (S_n - S_{n-1}) x^n$$

$$= \sum_{n=0}^{m} S_n x^n - \sum_{n=0}^{m} S_{n-1} x^n$$

$$= \sum_{n=0}^{m-1} S_n x^n + S_m x^m - x\sum_{n=0}^{m} S_{n-1} x^{n-1}$$

$$= \sum_{n=0}^{m-1} S_n x^n - x\sum_{n=0}^{m} S_{n-1} x^{n-1} + S_m x^m$$

$$= \sum_{n=0}^{m-1} S_n x^n - x\sum_{n=0}^{m-1} S_n x^n + S_m x^m \qquad [\because S_{-1} = 0]$$

$$= (1-x)\sum_{n=0}^{m-1} S_n x^n + S_m x^m.$$

Now, for $|x| < 1; x^m \to 0$ as $m \to \infty$ and $S_m \to S$.

$$\therefore \lim_{m\to\infty} \sum_{n=0}^{m} a_n x^n = \lim_{m\to\infty} (1-x)\sum_{n=0}^{m-1} S_n x^n + \lim_{m\to\infty} S_m x^m$$

$$\Rightarrow f(x) = \sum_{n=0}^{\infty} (1-x) S_n x^n \qquad (1)$$

Now, since $S_n \to S$, therefore for $\varepsilon > 0$, there exists integer N such that

$$|S_n - S| < \frac{\varepsilon}{2} \forall n \ge N \qquad (2)$$

Also, we have

$$(1-x)\sum_{n=0}^{\infty} x^n = 1 \qquad (3)$$

Hence for $n \ge N$, we have

$$|f(x) - S| = \left|(1-x)\sum_{n=0}^{\infty} S_n x^n - S\right|$$

$$= \left|(1-x)\sum_{n=0}^{\infty} S_n x^n - (1-x)\sum_{n=0}^{\infty} S x^n\right| \qquad \text{(by (3))}$$

$$= \left| (1-x) \sum_{n=0}^{\infty} (S_n - S) x^n \right|$$

$$\leq (1-x) \sum_{n=0}^{N} |S_n - S| x^n + \frac{\varepsilon}{2} (1-x) \sum_{n=N+1}^{\infty} x^n$$

$$\leq (1-x) \sum_{n=0}^{N} |S_n - S| x^n + \frac{\varepsilon}{2}$$

Now for a fixed N, $(1-x)\sum_{n=0}^{N} |S_n - S| x^n$ is continuous function of x having zero value at x = 1.

Thus, there exists $\delta > 0$ such that $1-\delta < x < 1$.

$$(1-x)\sum_{n=0}^{N} |S_n - S| x^n < \frac{\varepsilon}{2}$$

$\therefore \quad |f(x) - S| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$ whenever $1-\delta < x < 1$

Hence, $\quad \lim_{x \to 1^-} f(x) = S = \sum_{n=0}^{\infty} a_n$.

Remark 5. We state some result related to Cauchy product of two series which will use in following theorem, which is infact an application of Abel's theorem.

(i) Let $\sum_{n=0}^{\infty} a_n$ and $\sum_{n=0}^{\infty} b_n$, then the series $\sum_{n=0}^{\infty} c_n$ where $c_n = a_0 b_n + a_1 b_{n-1} + \dots\dots\dots + a_n b_0$ is called Cauchy product of series $\sum_{n=0}^{\infty} a_n$ & $\sum_{n=0}^{\infty} b_n$.

(ii) **Cauchy's Theorem.** Let $\sum_{n=0}^{\infty} a_n$ & $\sum_{n=0}^{\infty} b_n$ be absolutely convergent series such that $\sum a_n = A, \sum b_n = B$, then Cauchy's product series $\sum_n c_n$ is also absolutely convergent and $\sum_n c_n = AB$.

Theorem 9. If $\sum_{n=0}^{\infty} a_n$ & $\sum_{n=0}^{\infty} b_n$ and $\sum_{n=0}^{\infty} c_n$ converges to sum A, B & C respectively and if $\sum c_n$ be Cauchy product of $\sum a_n$ and $\sum b_n$ then AB = C.

Proof. $\sum_{n=0}^{\infty} c_n$ is the Cauchy product of $\sum_{n=0}^{\infty} a_n$ and $\sum_{n=0}^{\infty} b_n$.

$$\Rightarrow c_n = a_0 b_n + a_1 b_{n-1} + ...+ a_n b_0$$

Let $f(x)=\sum_{n=0}^{\infty} a_n x^n, g(x)=\sum_{n=0}^{\infty} b_n x^n$ and $h(x)=\sum_{n=0}^{\infty} c_n x^n; \forall 0 \le x \le 1$.

For $|x|<1$, the three series converge absolutely

$\therefore \quad \sum c_n x^n = f(x)g(x)$ (By Cauchy's theorem in Remark 5(ii))

$$\Rightarrow h(x)=f(x).g(x); 0\le x\le 1. \qquad (1)$$

Now by Abel's theorem

$$\lim_{x\to 1^-} f(x)=\sum_{n=0}^{\infty} a_n \Rightarrow f(x)\to A \text{ as } x\to 1^-$$

Similarly, $g(x)\to B,\ h(x)\to C$ as $x\to 1^-$ $\qquad (2)$

Thus from (1) & (2), we have

AB = C.

Example 1. Show that

$$(i) \tan^{-1} x = x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + ...$$

$$(ii) \frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + ...$$

Solution. (i) We know that

$$\left(1+x^2\right)^{-1} = 1 - x^2 + x^4 - x^6 + ...; |x|<1 \qquad (1)$$

The series on the right is a power series with radius of convergence 1, so it is absolutely convergent in (-1, 1) and uniformly convergent in [-k, k] where $|k| < 1$.

Now integrating (1), we get

$$\tan^{-1} x = c + x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + ...; |x| < 1$$

Putting x = 0, we obtain c = 0, so that

$$\tan^{-1} x = x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + ...; |x| < 1$$

The series on R.H.S is a power series with radius of convergence equal to 1. However, the series

$$x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + ...$$

is convergent at ± 1.

Hence by Abel's theorem, it is uniformly convergent in [-1, 1] and hence

$$\tan^{-1} x = x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + ...; -1 \le x \le 1$$

(ii) At x = 1. By Abel's theorem (Second form)

$$\tan^{-1} x = \lim_{x \to 1^-} \tan^{-1} x$$

$$\frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + ...$$

Example 2. Show that for $-1 \le x \le 1$,

$$(i) \log(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + ...$$

$$(ii) \log 2 = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + ...$$

Solution. (i) We know that

$$(1+x)^{-1} = 1 - x + x^2 - x^3 + ...; -1 \le x \le 1$$

On integrating, we get

$$\log(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - ....; -1 < x < 1$$

The power series on R.H.S. converges at x = 1.

So, by Abel's theorem

$$\log(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - ....; -1 \le x \le 1$$

(ii) Put x = 1, in above series we get result

$$\log(1+1) = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + ...; -1 \le x \le 1$$

$$\Rightarrow \log 2 = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + ...; -1 \le x \le 1.$$

Tauber's Theorem. The converse of Abel's theorem proved above is false in general. If f is given by

$$f(x) = \sum_{n=0}^{\infty} a_n x^n, \qquad -r < x < r$$

the limit $f(r-)$ may exists but yet the series $\sum_{n=0}^{\infty} a_n r^n$ may fail to converge. For example, if

$$a_n = (-1)^n, f(x) = \sum_{n=0}^{\infty} (-1)^n x^n$$

$$= \sum_{n=0}^{\infty} (-x)^n, |x| < 1$$

$$= 1 - x + x^2 - x^3 + \ldots\ldots\ldots\ldots$$

Then

$$f(x) = \frac{1}{1+x}, \quad -1 < x < 1.$$

$$f\left(1^-\right) = \lim_{x \to 1^-} f\left(x\right) \quad = \lim_{x \to 1^-} \frac{1}{1+x}.$$

Put x = 1- h, if $x \to 1^-$, $h \to 0$

$$= \lim_{h \to 0} \frac{1}{1+1-h} = \frac{1}{2}$$

$\Rightarrow f\left(1^-\right) exists.$

However,

$\sum_{n=0}^{\infty} a_n = \sum_{n=0}^{\infty} (-1)^n$ *is not convergent because this is oscillating between* -1 & 1.

Tauber showed that the converse of Abel's theorem can be obtained by imposing additional condition on coefficients a_n. A large number of such results are known now a days as Tauberian Theorems. We present here only Tauber's first theorem.

Theorem 10 (Tauber). Let $f(x) = \sum_{n=0}^{\infty} a_n x^n$, for $-1 < x < 1$ and suppose that $\lim_{n \to \infty} na_n = 0$. If $f(x) \to S$ as $x \to 1^-$, then $\sum_{n=0}^{\infty} a_n$ converges and has the sum S.

Proof. Let $n\sigma_n = \sum_{k=0}^{\infty} k\left|a_k\right|$. Then $\sigma_n \to 0$ *as* $n \to \infty$. (1)

Also, $\lim_{n \to \infty} f(x_n) = S$ *where* $x_n = 1 - \dfrac{1}{n}$. (2)

$$\left(\because \text{ when } n \to \infty, x_n \to 1^-, f(x_n) \to S\right).$$

Therefore to each $\in > 0$, we can choose an integer N such that $n \geq N$ implies

$$\left|\sigma_n - 0\right| < \frac{\in}{3}, \left|f(x_n) - S\right| < \frac{\in}{3}, \left|na_n - 0\right| < \frac{\in}{3}$$

i.e.,

$$\sigma_n < \frac{\in}{3}, \ |f(x_n) - S| < \frac{\in}{3}, \ n|a_n| < \frac{\in}{3} \quad \forall n \geq N. \qquad (3)$$

Let $S_n = \sum_{k=0}^{n} a_k$. Then for $-1 < x < 1$, we have

$$\begin{aligned}
S_n - S &= \sum_{k=0}^{n} a_k - S \\
&= \sum_{k=0}^{n} a_k - S + f(x) - \sum_{k=0}^{\infty} a_k x^k \\
&= f(x) - S + \sum_{k=0}^{n} a_k - \sum_{k=0}^{n} a_k x^k - \sum_{k=n+1}^{\infty} a_k x^k \\
&= f(x) - S + \sum_{k=0}^{n} a_k (1 - x^k) - \sum_{k=n+1}^{\infty} a_k x^k .
\end{aligned}$$

$$|S_n - S| = \left| f(x) - S + \sum_{k=0}^{n} a_k (1 - x^k) - \sum_{k=n+1}^{\infty} a_k x^k \right|. \qquad (4)$$

Let $x \in (0,1)$. Then

$$(1 - x^k) = (1 - x)(1 + x + \ldots\ldots. + x^{k-1}) \leq k(1 - x)$$

for each k. Therefore, if $n \geq N$ and $0 < x < 1$, we have

$$\begin{aligned}
|S_n - S| &= \left| f(x) - S + \sum_{k=0}^{n} a_k (1 - x^k) - \sum_{k=n+1}^{\infty} a_k x^k \right| \\
&\leq |f(x) - S| + \left| \sum_{k=0}^{n} a_k (1 - x^k) \right| + \left| \sum_{k=n+1}^{\infty} a_k x^k \right| \\
&< \ |f(x) - S| + \left| \sum_{k=0}^{n} a_k k(1 - x) \right| + \left| \sum_{k=n+1}^{\infty} a_k x^k \right| \\
&\leq |f(x) - S| + (1 - x) \sum_{k=0}^{n} k|a_k| + \left| \sum_{k=n+1}^{\infty} a_k x^k \right| \\
&< |f(x) - S| + (1 - x) \sum_{k=0}^{n} k|a_k| + \frac{\in}{3n} \sum_{k=n+1}^{\infty} x^k \\
&< |f(x) - S| + (1 - x) \sum_{k=0}^{n} k|a_k| + \frac{\in}{3n(1 - x)} .
\end{aligned}$$

Putting $x = x_n = 1 - \frac{1}{n}$, we find that

$$\Rightarrow \ (1 - x) = \frac{1}{n}$$

$$|S_n - S| < \; |f(x) - S| + \sum_{k=0}^{n} \frac{k|a_k|}{n} + \frac{\in}{3n.\frac{1}{n}}$$

$$< \frac{\in}{3} + \frac{\in}{3} + \frac{\in}{3} = \in.$$

$\sum_{n=0}^{\infty} a_k$ converges & has sum S, which completes the proof.

3.3 Functions of Several Variables

This section is devoted to calculus of functions of several variables in which we study derivatives and partial derivatives of functions of several variables along with their properties. The notation for a function of two or more variables is similar to that for a function of a single variable. A function of two variables is a rule that assigns a real number f(x, y) to each pair of real numbers (x, y) in the domain of the function which can be extended to three and more variables.

3.3.1 Linear transformation

Definition 1. A mapping f of a vector space X into a vector space Y is said to be a linear transformation if

$$\begin{aligned} & f(x_1 + x_2) = f(x_1) + f(x_2), \\ & f(cx) = cf(x) \end{aligned}$$

for all $x, x_1, x_2 \in X$ and all scalars c.

Clearly, if f is linear transformation, then f(0) = 0.

A linear transformation of a vector space X into X is called linear operator on X.

If a linear operator T on a vector space X is one-to-one and onto, then T is invertible and its inverse is denoted by T^{-1}. Clearly, $T^{-1}(Tx) = x$ for all $x \in X$. Also, if T is linear, then T^{-1} is also linear.

Theorem 1. A linear operator T on a finite dimensional vector space X is one-to-one if and only if the range of T is equal to X. i.e, T(X) = X.

Proof. Let R(T) denotes range of T. Let $\{x_1, x_2, ..., x_n\}$ be basis of X. Since T is linear the set $\{Tx_1, Tx_2, ..., Tx_n\}$ spans R(T). The range of T will be whole of X if and only if $\{Tx_1, Tx_2, ..., Tx_n\}$ is linearly independent.

So, suppose first that T is one-to-one. We shall prove that $\{Tx_1, Tx_2, ..., Tx_n\}$ is linearly independent. Hence, let

$$c_1 Tx_1 + c_2 Tx_2 + ... + c_n Tx_n = 0$$

Since T is linear, this yields

$$T(c_1x_1 + c_2x_2 + ... + c_nx_n) = 0$$

and so $$c_1x_1 + c_2x_2 + ... + c_nx_n = 0$$

Since $\{x_1, x_2, ..., x_n\}$ is linearly independent, we have, $c_1 = c_2 = ...c_n = 0$.

Thus $\{Tx_1, Tx_2, ..., Tx_n\}$ is linearly independent and so R(T) = X if T is one-to-one.

Conversely, suppose that $\{Tx_1, Tx_2, ..., Tx_n\}$ is linearly independent and so

$$c_1Tx_1 + c_2Tx_2 + ... + c_nTx_n = 0 \tag{1}$$

implies $c_1 = c_2 = ...c_n = 0$. Since T is linear (1) implies

$$T(c_1x_1 + c_2x_2 + ... + c_nx_n) = 0$$

$$\Rightarrow \quad c_1x_1 + c_2x_2 + ... + c_nx_n = 0$$

Thus T(x) = 0 only if x = 0. Now

$$T(x) = T(y) \quad \Rightarrow \; T(x-y) = 0 \quad \Rightarrow \; x - y = 0 \quad \Rightarrow \; x = y$$

and so T is one-to-one. This completes the proof of theorem.

Definition 2. Let L(X, Y) be the set of all linear transformations of the vector space X into the vector space Y. If $T_1, T_2 \in L(X,Y)$ and if c_1, c_2 are scalars, then

$(c_1T_1+c_2T_2)(x)= c_1T_1x+c_2 T_2x$; $x \in X$. It can be shown that $c_1T_1+c_2T_2 \in L(X, Y)$.

Definition 3. Let X, Y and Z be vector spaces over the same field. If S, T $\in$ L(X,Y), then we define their product ST by

$$ST(x) = S(T(x));\ x \in X.$$

Also, $ST \in L(X, Y)$.

Euclidean space $\mathbf{R^n}$. A point in two dimensional space is an ordered pair of real no. (x_1, x_2). Similarly, a point in three dimensional space is an ordered triplet of real no. (x_1, x_2, x_3). It is just as easy to consider an ordered n-tuple of real no. $(x_1, x_2, \ldots, x_n)$ and refer to this as a point in n-dimensional space.

Definition 4. Let n > 0 be an integer. An ordered set of n real no. $(x_1, x_2, \ldots, x_n)$ is called an n-dimensional point or a vector with n-component points. Vector will usually be denoted by single bold face letter.

e.g. $x = (x_1, x_2, \ldots, x_n)$

$y = (y_1, y_2, \ldots, y_n)$

The number x_k is called the k^{th} co-ordinate of point x or k^{th} component of the vector x.

The set of all n-dimensional point is called n-dimensional Euclidean space or n-space and is denoted by $\mathbf{R^n}$.

Algebraic operations in R^n - n-dimensional Euclidean space are as follow:

Let $x = (x_1,x_2,\ldots\ldots,x_n)$ and $y = (y_1,y_2,\ldots\ldots\ldots,y_n)$ be in R^n.

We define

(a) Equality $x = y$ iff $x_1 = y_1$, $x_2 = y_2$, …………, $x_n = y_n$.

(b) Sum $x + y = ((x_1 + y_1, x_2 + y_2,\ldots\ldots,x_n + y_n)$

(c) Multiplication by real no. (Scalar):

$ax = a(x_1, x_2,\ldots\ldots, x_n) = (ax_1,ax_2,\ldots\ldots, ax_n)$

(d) Difference $x - y = x+(-)y$

(e) Zero vector or origin $0 = (0, 0, \ldots\ldots ,0)$.

(f) Inner product or dot product

$$xy = \sum_{k=1}^{n} x_k y_k .$$

(g) For all $x \in R^n$. Also if λ is such that

$$|Tx| \le \lambda |x|,\ x \in R^n,\ \text{then} \|T\| \le \lambda.$$

(h) Norm or length

If $T \in L(R^n, R^m)$. Then

$$\text{lub}\{|Tx| : x \in R^n, |x| \le 1\}$$

is called Norm of T and is denoted by $||T||$. The inequality

$$\|Tx\| \le \|T\|\|x\|$$

and

$$\|x\| = \left(\sum_{k=1}^{n} x_k^2\right)^{1/2} .$$

The norm $\|x-y\|$ is called the distance between x & y.

(i) Also, Let x and y denote points in R^n, then the following results hold:

(i) $\|x\| \ge 0$ *and* $\|x\| = 0$ *iff* $x = 0$.

(ii) $\|ax\| = |a|.\|x\|$ for every real a.

(iii) $\|x-y\| = \|y-x\|$

(iv) **Cauchy Schwarz Inequality:**

$$|<x.y>|^2 \le \|x\|\|y\|.$$

(v) $\|x+y\| \le \|x\| + \|y\|.$

Note 1. Sometimes the triangle inequality is written in the form

$$\|x-z\| \le \|x-y\| + \|y-z\|.$$

This follows form (v) by replacing x by x-y and y by y-z. We also have

$$\|\|x\| - \|y\|\| \le \|x-y\|$$

Definition 5. The unit co-ordinate vector u_k in R^n is the vector whose k^{th} component is 1 and remaining components are zero. Then

$$u_1 = (1,\ 0,\ 0,\ \ldots\ldots,\ 0)$$
$$u_2 = (0,\ 1,\ 0, \ldots\ldots, 0)$$
$$\ldots\ldots\ldots\ldots\ldots\ldots$$
$$\ldots\ldots\ldots\ldots\ldots\ldots$$
$$u_n = (0,\ 0,\ \ldots\ldots\ldots\ldots, 1).$$

$If\ \ x = (x_1, x_2, \ldots\ldots\ldots\ldots, x_n),\ then$

$$x = x_1u_1 + x_2u_2 + \ldots\ldots\ldots + x_nu_n$$
$$\&$$
$$x_1 = xu_1, x_2 = xu_2, \ldots\ldots\ldots\ldots, x_n = xu_n.$$

The vectors $u_1, u_2, \ldots\ldots\ldots, u_n$ are also called basis vectors.

Theorem 2. Let $T, S \in L(R^n, R^m)$ and c be a scalar. Then

$(a)\|T\| < \infty$ *and T is uniformly continuous mappings of* R^n *and* R^m.

$(b)\|T+S\| \le \|T\| + \|S\|$ *and* $\|cT\| = |c|\|T\|$.

(c) *If* $d(T,S) = \|T-S\|$, *then d is a metric.*

Proof. (a) Let $\{e_1, e_2, \ldots, e_n\}$ be the standard basis in R^n and let $x \in R^n$. Then $x = \sum_{i=1}^{n} c_i e_i$.

Suppose $|x| < 1$ so that $|c_i| \le 1$ for $i = 1, 2, \ldots, n$. Then

$$|Tx| = \left|\sum c_i Te_i\right| \le \sum |c_i|.|Te_i|$$
$$\le \sum |Te_i|$$

Taking lub over $x \in R^n, |x| \leq 1$

$$\|Tx\| \leq \sum |Te_i| < \infty.$$

Further

$\|Tx - Ty\| = \|T(x-y)\| \leq \|T\|\|x-y\|$; $x, y \in R^n$

So if $\|x-y\| < \dfrac{\in}{\|T\|}$, *then*

$\|Tx - Ty\| < \in;\ x, y \in R^n.$

Hence, T is uniformly continuous.

(b) We have

$$\begin{aligned}|(\mathrm{T}+\mathrm{S})\mathrm{x}| &= |\mathrm{Tx}+\mathrm{Sy}| \\ &\leq |\mathrm{Tx}| + |\mathrm{Sx}| \\ &\leq \|\mathrm{T}\||\mathrm{x}| + \|\mathrm{S}\||\mathrm{x}| \\ &= (\|\mathrm{T}\| + \|\mathrm{S}\|)|\mathrm{x}|\end{aligned}$$

Taking lub over $x \in R^n, |x| \leq 1,$ *we have*

$$\|T+S\| \leq \|T\| + \|S\|.$$

Similarly, it can be shown that

$$\|\mathrm{cT}\| = |\mathrm{c}|\|\mathrm{T}\|.$$

(c) We have $\mathrm{d(T,S)} = \|\mathrm{T-S}\| \geq 0$ and $\mathrm{d(T,S)} = \|\mathrm{T-S}\| = 0 \Leftrightarrow \mathrm{T=S}.$

Also $\mathrm{d(T,S)} = \|\mathrm{T-S}\| = \|\mathrm{S-T}\| = \mathrm{d(S,T)}$

Further, if $\mathrm{S,T,U} \in \mathrm{L}(\mathrm{R}^n, \mathrm{R}^m)$, then

$$\begin{aligned}\|\mathrm{S-U}\| &= \|\mathrm{S-T+T-U}\| \\ &\leq \|\mathrm{S-T}\| + \|\mathrm{T-U}\|\end{aligned}$$

Hence, d is a metric.

Theorem 3. If T ϵ L(R^n, R^m) and S ϵ L(R^n, R^m), then

$$\|\mathrm{ST}\| \leq \|\mathrm{S}\|\|\mathrm{T}\|$$

Proof. We have

$$\begin{aligned}|(\mathrm{ST})\mathrm{x}| = |\mathrm{S(Tx)}| &\leq \|\mathrm{S}\||\mathrm{Tx}| \\ &\leq \|\mathrm{S}\|\|\mathrm{T}\||\mathrm{x}|\end{aligned}$$

Taking sup *over* x, $|x| \le 1$, *we have*

$$\|ST\| \le \|S\|\|T\|.$$

In theorem 2, we have seen that the set of linear transformation form a metric space. Hence the concepts of convergence, continuity, open sets etc. make sense in R^n.

Theorem 4. Let C be the collection of all invertible linear operators on R^n.

(a) If $T \in C, \|T^{-1}\| = \frac{1}{\alpha}$, $S \in L(R^n, R^m)$ and $\|S - T\| = \beta < \alpha$, then $S \in C$.

(b) C is an open subset of $L(R^n, R^m)$ and mapping $T \to T^{-1}$ is continuous on C.

Proof. We note that

$$\begin{aligned} |x| = |T^{-1}Tx| &\le \|T^{-1}\||Tx| \\ &\le \frac{1}{\alpha}|Tx| \text{ for all } x \in R^n \end{aligned}$$

and so

$$\begin{aligned} (\alpha - \beta)|x| &= \alpha|x| - \beta|x| \\ &\le |Tx| - \beta|x| \\ &\le |Tx| - |(S-T)x| \\ &\le |Sx| \qquad \forall\ x \in R^n. \end{aligned} \tag{1}$$

Thus kernel of S consists of 0 only. Hence S is one-to-one. Then Theorem 1 implies that S is also onto. Hence S is invertible and so S ϵ C. But this holds for all S satisfying ||S - T|| < α. Hence every point of C is an interior point and so C is open.

Replacing x by $S^{-1}y$ in (1), we have

$$(\alpha - \beta)|S^{-1}y| \le |SS^{-1}y| = |y|$$

or

$$|S^{-1}y| \le \frac{|y|}{\alpha - \beta}$$

and so

$$\|S^{-1}\| \le \frac{1}{\alpha - \beta}$$

since

$$S^{-1} - T^{-1} = S^{-1}(T - S)T^{-1}$$

We have

$$\begin{aligned} \|S^{-1} - T^{-1}\| &= \|S^{-1}\|\|T - S\|\|T^{-1}\| \\ &\le \frac{\beta}{\alpha(\alpha - \beta)}. \end{aligned} \tag{2}$$

Thus if f is the mapping which maps $T \to T^{-1}$, then (2) implies

$$\|f(S)-f(T)\| \le \frac{\|S-T\|}{\alpha(\alpha-\beta)}.$$

Hence, if $\|S - T\| \to 0$ then $f(S) \to f(T)$ and so f is continuous. This completes the proof of the theorem.

3.3.2 Derivatives in an open subset E of $\mathbf{R}^n$

In one-dimensional case, a function f with a derivative at c can be approximated by a linear polynomial. In fact if f′(c) exists, let r(h) denotes the difference

$$r(h)=\frac{f(x+h)-f(x)}{h}-f'(x) \quad if \quad h\neq 0 \qquad (1)$$

and let r(0) = 0. Then we have

$$f(x+h)=f(x)+h\,f'(x)+hr(h), \qquad (2)$$

an equation which holds also for h = 0. The equation (2) is called the **First order Taylor formula for approximating** f(x + h) – f(x) by h f′(x). The error committed in this approximation is h r(h). From (1), we observe that $r(h) \to 0$ as $h \to 0$. The error h r(h) is said to be of smaller order than h as $h \to 0$. We also note that h f′(x) is a linear function of h. Thus, if we write Ah = h f′(x), then

$$A(ah_1+bh_2)=aAh_1+bAh_2$$

Here, the aim is to study total derivative of a function **f** from $\mathbf{R}^n$ to $\mathbf{R}^m$ in such a way that the above said properties of hf′(x) and hr(h) are preserved.

Definition 1(Open ball and open sets in $\mathbf{R}^n$). Let 'a' be a given point in R^n and let r be a given positive number, then the set of all points x in R^n such that

$\|x-a\|<r$ is called an open n-ball of radius 'r' and centre 'a'.

We denote this set by B(a) or B(a, r) . The B (a, r) consists of all points whose distance from 'a' is less than r.

In R^1, this is simply an open interval with centre at a.

In R^2, it is a circular disc.

In R^3, it is a spherical solid with centre at a and radius r.

Definition 2 (Interior point). Let E be a subset of R^n and assume that $a \in E$, then a is called an interior point of E if there is an open ball with centre surrounded by an n-ball. i.e.,

$$B(a)\subseteq E.$$

The set of all interior points of E, is called the interior of E and is denoted by int E.

Any set containing a ball with centre 'a' is sometime called a neighbourhood of a.

Definition 3 (Open set). A set E in R^n is called open if all points are interior points.

Note 1. A set E is open if and only if E = interior of E.

Every open n-ball is an open set in R^n.

The cartesian product $(a_1, b_1) \times (a_2, b_2) \times .... \times (a_n, b_n)$ of n-dimensional open interval $(a_1,b_1),\ldots\ldots\ldots,(a_n,b_n)$ is an open set is R^n called n-dimensional open interval, we denote it by (a, b) where

$$a = (a_1, a_2, ..., a_n)$$
$$b = (b_1, b_2, ..., b_n).$$

Remark 1. (i) Union of any collection of open sets is an open set.

(ii) The intersection of a finite collection of open sets is open.

(iii) Arbitrary intersection of open sets need not be open.

e.g. Consider the seq. of open interval such that

$$G_n = \left\{ -\frac{1}{n}, \frac{1}{n} \right\}; \ n \in N$$

Clearly each G^n is open set but $G_1 \cap G_2 .... G_n = \{0\}$, which is being a finite set is not open.

Definition 4 (The structure of open sets in R'). In R' the union of countable collection of disjoint open interval is an open set in R' can be obtained in this way.

First we introduce the concept of a component interval.

Definition 5 (Component interval). Let E be an open subset in R' and open interval I (which may be finite or infinite) is called a component interval of E

$$If \ I \subseteq E \ and \ if \ there \ is \ no \ interval \ J \neq I \ s.t. \ I \subseteq J \subseteq E.$$

In other words, a component interval of E is not a proper subset of any other open interval contained in E.

Remark 2. (i) Every point of a nonempty open set E belongs to one and only one component interval of E.

(ii) Representative theorem for open sets on the real line.

Every nonempty open set E in R' is the union of a countable collection of disjoint intervals of E.

Definition 6(Closed set). A set in R^n is called closed if and only if its complement R^n - E is open.

Remark 3. (i) The union of a finite collection of closed sets is closed and the intersection of an arbitrary collection of closed set is closed.

(ii) If A is open and B is closed, then A - B is open and B-A is closed.

$$A - B = A \cap B^c.$$

Definition 7 (Adherent point). Let E be a subset of R^n and x is point in R^n, x is not necessary in E. Then x is said to be adherent to E if every n-ball B(x) contains atleast one point of E.

E.g. (i) If x϶E, then x adherenes to E for the trivial reason that every n-ball B(x) contains x.

(ii) If E is a subset of R which is bounded above. Then sup.E is adherent to E.

Some points adheres to E because every ball B(x) contains points of E distinct from x these are called adherent points.

Definition 8 (Accumulation point/Limit point). Let E be a subset of R^n and x is a point in R^n, then x is called an accumulation point of E if every n-ball B(x) contains atleast one point of E distinct from x.

In other words, x is an accumulation point of E if and only if x adheres to E-{x}.

If x ϵ E, but x is not an accumulation point of E, then x is called an isolated point of E.

e.g. (i) The set of numbers of the form 1/n (n=1,2,……..) has 0 as an accumulation point.

(ii) The set of rational numbers has every real number as accumulation point.

(iii) Every point of the closed interval [a, b] is an accumulation point of the set of numbers in the open interval (a, b).

Remark 4. If x is an accumulation point of E, then every n-ball B(x) contains infinitely many points of E.

Definition 9 (Closure of a set). The set of all adherent points of a set E is called a closure of E *and is denoted by* $\overline{E}$.

Definition 10 (Derived set). The set of all accumulation points of a set E is called the derived set of E and is denoted by E'.

Remark 5. (i) A set E in R^n is closed if and only if it contains all its adherent points.

(*ii*) *A set E is closed iff* $E = \overline{E}$.

(iii) A set E in R^n is closed iff it contains all its accumulation points.

Definition 11. Suppose E is an open set in R^n and let $f : E \to R^n$ be a function defined on a set E in R^n with values in R^m. Let x ϵ E and h be a point in R^n such that $|h| < r$ and x + h ϵ B(x, r). Then f is said to be differentiable at x if there exists a linear transformation A of R^n into R^n such that

$$f(x+h)=f(x)+Ah+r(h) \qquad (1)$$

where the reminder r(h) is small in the sense that

$$\lim_{h\to 0} \frac{|r(h)|}{|h|} = 0.$$

We write f′(x) = A.

The equation (1) is called a **First order Taylor formula.**

$$h\to 0 \frac{|f(x+h)-f(x)-Ah|}{|h|}=0. \qquad (2)$$

The equation (2) thus can be interpreted as "For fixed x and small h, f(x + h) – f(x) is approximately equal to f′(x)h, that is, the value of a linear function applied to h."

Also (1) shows that f is continuous at any point at which f is differentiable.

The derivatives Ah derived by (1) or (2) is called **total derivative** of **f** at **x** or the **differential of f at x.**

In particular, let f be a real valued function of three variables x, y, z say. Then f is differentiable at the point (x, y, z) if it possesses a determinant value in the neighbourhood of this point and if

$\Delta f = f(x+\Delta x, y+\Delta y, z+\Delta z)-f(x,y,z) = A\Delta x + B\Delta y + C\Delta z + \in \rho$, where $\rho = |\Delta x| + |\Delta y| + |\Delta z|$, $\in \to 0$ as $\rho \to 0$ and A, B, C are independent of x, y, z. In this case $A\Delta x + B\Delta y + C\Delta z$ is called differential of f at (x, y, z).

Theorem 1 (Uniqueness of derivative of a function). Let E be an open set in R^n and f maps E in R^m and x ϵ E. Suppose h ϵ R^n is small enough such that x + h ϵ E. Then f has a unique derivative.

Proof. If possible, let there are two derivatives A_1 and A_2. Therefore

$$\lim_{h\to 0} \frac{|f(x+h)-f(x)-A_1 h|}{|h|}=0$$

and

$$\lim_{h\to 0} \frac{|f(x+h)-f(x)-A_2 h|}{|h|}=0$$

Consider B = A_1 - A_2. Then

$$\begin{aligned} Bh &= A_1h - A_2h \\ &= f(x + h)- f(x) +f(x)- f(x + h) + A_1h - A_2h \\ &= f(x + h)- f(x) - A_2h + f(x)- f(x + h) + A_1h \end{aligned}$$

and so

$$|Bh| < |f(x+h)-f(x)-A_2h| + |f(x+h)-f(x)-A_1h|$$

which implies

$$\lim_{h\to 0}\frac{|Bh|}{|h|} \le \lim_{h\to 0}\frac{|f(x+h)-f(x)-A_1h|}{|h|} + \frac{|f(x+h)-f(x)-A_2h|}{|h|}$$
$$=0$$

For fixed h ≠ 0, it follows that

$$\frac{|B(th)|}{|th|} \to 0 \to as\ t \to 0. \qquad (1)$$

The linearity of B shows that L.H.S of (1) is independent of t. Thus Bh = 0 for all $h \in R^n$. Hence B = 0, that is, $A_1 = A_2$, which proves uniqueness of the derivative.

The following theorem, known as chain rule, tells us how to compute the total derivatives of the composition of two functions.

Theorem 2 (Chain rule). Suppose E is an open set in R^n, f maps E into R^m, f is differentiable at x_0 with total derivative $f'(x_0)$, g maps an open set containing f(E) into R^k and g is differentiable at $f(x_0)$ with total derivative $g'(f(x_0))$. Then the composition map F = fog, a mapping E into R^k and defined by F(x) = g(f(x)) is differentiable at x_0 and has the derivative

$$F'(x_0) = g'(f(x_0))\, f'(x_0).$$

Proof. Take

$$y_0 = \mathbf{f}(x_0),\ A = \mathbf{f'}(x_0),\ B = \mathbf{g'}(y_0)$$

and define

$$\begin{aligned} r_1(x) &= f(x) - f(x_0) - A(x - x_0) \\ r_2(y) &= g(y) - g(y_0) - B(y - y_0) \\ r(x) &= F(x) - F(x_0) - BA(x - x_0). \end{aligned}$$

To prove the theorem, it is sufficient to show that

$$F'(x_0) = BA,$$

that is,

$$\frac{r(x)}{|x - x_0|} \to 0 \quad as\, x \to x_0 \tag{1}$$

But, in term of definition of F(**x**), we have

$$r(x) = g(f(x)) - g(y_0) - B(f(x) - f(x_0) - A(x - x_0))$$

so that

$$r(x) = r_2(f(x)) + B\, r_1(x). \tag{2}$$

If $\epsilon > 0$, it follows from the definitions of A and B that there exists $\eta > 0$ and $\delta > 0$ such that

$$\frac{|r_2(y)|}{|y - y_0|} \le \in \quad as\ y \to y_0$$

$$or\quad |r_2(y)| \le \in |y - y_0| \quad as\, |y - y_0| < \eta \quad i.e.,\ |f(x) - f(x_0)| < \eta$$

$$and\, |r_1(x)| \le \in |x - x_0| \quad if\quad |x - x_0| < \delta.$$

Hence

$$\begin{aligned} |r_2(f(x))| &\le \in |f(x) - f(x_0)| \\ &= \in |r_1(x) + A(x - x_0)| \\ &\le \in^2 |x - x_0| + \in \|A\|(x - x_0) \end{aligned} \tag{3}$$

and

$$|Br_1(x)| \le \|B\| |r_1(x)|$$
$$\le \in \|B\| |x - x_0| \text{ if } |x - x_0| < \delta. \qquad (4)$$

Using (3) and (4), the expression (2) yields

Hence

$$|r(x)| \le \in^2 |x - x_0| + \in \|A\|(x - x_0) + \in \|B\|(x - x_0)$$
$$\frac{|r(x)|}{|x - x_0|} \le \in^2 + \in \|A\| + \in \|B\|$$
$$= \in [\in + \|A\| + \|B\|] \text{ if } |x - x_0| < \delta$$

Hence,

$$\frac{|r(x)|}{|x - x_0|} \to 0 \text{ as } x \to x_0$$

which in turn implies

$$F'(x_0) = BA = g'(f(x_0))\, f'(x_0).$$

3.3.3 Partial derivatives.

Let $\{e_1, e_2, ... e_n\}$ be the standard basis of R^n. Suppose f maps an open set $E \subset R^n$ into R^m and let $f_1, f_2, ..., f_m$ be components of f. Define $D_k f_i$ on E by

$$(D_k f_i)(x) = \lim_{t \to 0} \frac{f_i(x + te_k) - f_i(x)}{t} \qquad (1)$$

provided the limit exists.

Writing $f_i(x_1, x_2, ..., x_n)$ in place of $f_i(x)$ we observe that $D_k f_i$ is derivative of f_i with respect to x_k, keeping the other variable fixed. That is why, we use $\frac{\partial f_i}{\partial x_k}$ frequently in place of $D_k f_i$.

Since $f = (f_1, f_2, ..., f_n)$, we have

$$D_k f(x) = (D_k f_1(x), D_k f_2(x), ..., D_k f_n(x))$$

which is partial derivative of **f** with respect to x_k.

Furthermore, if **f** is differentiable at **x**, then the definition of f′(**x**) shows that

$$\lim_{t \to 0} \frac{f(x + th_k) - f(x)}{t} = f'(x) h_k \qquad (2)$$

If we take $h_k = e_k$, taking components of vector in (2), it follows that

"If f is differentiable at x then all partial derivatives $D_k f_i$ (x) exist".

In particular, if f is real valued (m=1), then (1) takes the form

$$(D_k f)(x) = \lim_{t\to 0} \frac{f(x+t)-f(x)}{t}.$$

For example, if f is a function of three variables x, y and z, then

$$Df(x) = \lim_{\Delta x\to 0} \frac{f(x+\Delta x, y, z)-f(x,y,z)}{\Delta x}$$

$$Df(y) = \lim_{\Delta y\to 0} \frac{f(x, y+\Delta y, z)-f(x,y,z)}{\Delta y}$$

and

$$Df(z) = \lim_{\Delta z\to 0} \frac{f(x, y, z+\Delta z)-f(x,y,z)}{\Delta z}$$

and are known respectively as partial derivatives of f with respect to x, y, z.

The next theorem shows that $A\,h = f'(\mathbf{x})(h)$ is a linear combination of partial derivatives of f.

Theorem 1. Let $E \subseteq \mathbf{R}^n$ and let $f : E \to \mathbf{R}^n$ be differentiable at **x** (interior point of open set E). If $h = c_1e_1 + c_2e_2 + \ldots + c_ne_n$ where $\{e_1, e_2, \ldots, e_n\}$ is a standard basis for $\mathbf{R}^n$, then

$$f'(\mathbf{x})(h) = \sum_{k=1}^{n} c_k D_k f(x).$$

Proof. Using the linearity of $f'(\mathbf{x})$, we have

$$f'(\mathbf{x})(h) = \sum_{k=1}^{n} f'(x)(c_k e_k)$$

$$= \sum_{k=1}^{n} c_k f'(x) e_k$$

But, by (2),

$$f'(\mathbf{x})e_k = (D_k f)(x)$$

Hence

$$f'(\mathbf{x})(h) = \sum_{k=1}^{n} c_k D_k (f)(x)$$

If f is real valued (m = 1), we have

$$f'(\mathbf{x})(h) = (D_1 f(x), D_2 f(x), \ldots, D_n f(x))\,h.$$

Definition 1 (Continuously differentiable mapping).

A differentiable mapping f of an open set $E \subset R^n$ into R^m is said to be continuously differentiable in E if f′ is continuous mapping of E into $L(R^n, R^m)$.

Thus to every $\epsilon > 0$ and every $x \in E$ there exists a $\delta > 0$ such that
$\|f'(y) - f'(x)\| < \in$ *if* $y \in E$ *and* $|y - x| < \delta$.

In this case we say that **f** is a C′-mapping in E or that **f** $\in$ C′(E).

Theorem 2. Suppose f maps an open set $E \subset R^n$ into R^m. Then f is continuously differentiable if and only if the partial derivatives D_jf_i exist and are continuous on E for $1 \le i \le m, 1 \le j \le n$.

Proof. Suppose first that f is continuously differentiable in E. Therefore to each $x \in E$ and $\epsilon > 0$, there exists a $\delta > 0$ such that

$$\|f'(y) - f'(x)\| < \in \text{ if } y \in E \text{ and } |y - x| < \delta.$$

We have then

$$\begin{aligned} |f'(y)e_j - f'(x)e_j| &= |(f'(y) - f'(x))e_j| \\ &\le \|f'(y) - f'(x)\|\|e_j\| \\ &= \|f'(y) - f'(x)\| \quad < \in \text{ if } y \in E \text{ and } |y - x| < \delta. \end{aligned} \tag{1}$$

Since f is differentiable, partial derivatives D_jf_i exist. Taking components of vectors in (1), it follows that

$$|(D_jf_i)(y) - D_jf_i(x)| < \in \text{ if } y \in E \text{ and } |y - x| < \delta.$$

Hence D_jf_i are continuous on E for $1 \le i \le m, 1 \le j \le n$.

Conversely, suppose that D_jf_i are continuous on E for $1 \le i \le m, 1 \le j \le n$. It is sufficient to consider one-dimensional case, i.e., the case m = 1. Fix $x \in E$ and $\epsilon > 0$. Since E is open, x is an interior point of E and so there is an open ball $B \subset E$ with centre at x and radius R. The continuity of D_jf implies that R can be chosen so that

$$|(D_jf)(y) - (D_jf)(x)| < \frac{\in}{n} \text{ if } y \in B, 1 \le j \le n. \tag{2}$$

Suppose $h = \sum h_je_j, |h| < R$, and take $v_0 = 0$

and $$v_k = h_1e_1 + h_2e_2 + \dots + h_ke_k \text{ for } 1 \le k \le n.$$

Then

$$f(x+h) - f(x) = \sum_{j=1}^{n} [f(x + v_j) - f(x + v_{j-1})]. \tag{3}$$

$$\sum_{j=1}^{n} \left| f\left(s + v_{j-i} + h_je_j\right) - f(x + v_{j-1}) \right|$$

Mean value theorem implies

$$f(x+h)-f(x)=\sum_{j=1}^{n}\left|h_j D_j\, f\left(x+v_{j-1}+\theta_j h_j e_j\right)\right| \textit{for some } \theta \in (0,1)$$

subtracting $\sum_{j=1}^{n} h_j \left(D_j f\right)(x)$

and then taking modulus,

$$\left|f(x+h)-f(x)-\sum_{j=1}^{n} h_j\left(D_j f\right)(x)\right|=\left|\sum_{j=1}^{n}\left[h_j D_j f\left(x+v_{j-1}+\theta_j h_j e_j\right)-h_j\left(D_j f\right)(x)\right]\right|$$

$$=\left|\sum_{j=1}^{n}\left[h_j\left[(D_j f)\left(x+v_{j-1}+\theta_j h_j e_j\right)\right]-\left(D_j f\right)(x)\right]\right|$$

$$<\sum_{j=1}^{n}|h_j|\frac{\varepsilon}{n}$$

$$<\frac{\varepsilon}{n}\sum_{j=1}^{n}|h|=\frac{\varepsilon}{n}.n\,|h|=\varepsilon\,|h|$$

Hence f is differentiable at x and f'(x) is the linear function which assigns the number

$f'(x)h=\sum_{j=1}^{\infty} h_j\left(D_j f\right)(x)$ when f'(x) is applied on h. Since $\left(D_1 f\right)(x),\left(D_2 f\right)(x)...,\left(D_n f\right)(x)$ are continuous functions on E, it follows that f' is continuous and hence $f \in C'(E)$.

Hence f is differentiable at x and f′(x) is the linear function which assigns the number $\sum h_j(D_i f)(x)$ to the vector $h=\sum h_j e_j$. The matrix [f′(x)] consists of the row $((D_1 f)(x),(D_2 f)(x)....,(D_n f)(x))$. Since $(D_1 f)(x),(D_2 f)(x)....,(D_n f)(x)$ are continuous functions on E, it follows that f′ is continuous and hence $f \in C'(E)$.

Classical theory for functions of more than one variable

Consider a variable u connected with the three independent variables x, y and z by the functional relation

$$u = u(x, y, z)$$

If arbitrary increment Δx, Δy, Δz are given to the independent variables, the corresponding increment Δu of the dependent variable of course depends upon three increments assigned to x, y, z.

Definition 2 (Continuous function). Let $u : R^n \to R$ be a function. Then u is said to be continuous at a point $x=\left(x_1, x_2, ..., x_n\right) \in R^n$. If given $\varepsilon > 0$, there exists a $\delta > 0$ such that

$$\left|u\left(x_1, x_2, ..., x_n\right)-u\left(x_1+\Delta x_1, x_2+\Delta x_2, ..., x_n+\Delta x_n\right)\right| < \varepsilon$$

whenever $\rho=\sqrt{\Delta x_1^2+\Delta x_2^2+............+\Delta x_n^2}<\delta$.

Definition 3 (Differentiable function). A function u = u(x, y, z) is said to be differentiable at point (x, y, z) if it posses a determinant value in the neighbourhood of this point and if

$$\Delta u = A\Delta x + B\Delta y + C\Delta z + \in \rho,$$

where $\rho = |\Delta x| + |\Delta y| + |\Delta z|, \in \to 0$ as $\rho \to 0$ and A, B, C are independent of $\Delta x, \Delta y, \Delta z$.

In the above definition ρ may always be replaced by η, where

$$\eta = \sqrt{\Delta x^2 + \Delta y^2 + \Delta z^2}\ .$$

So, if $u : R^n \to R$ be a function, then u is said to be differentiable at a point $x = (x_1, x_2, \ldots\ldots\ldots, x_n) \in R^n$ if there exist constants $A_1, A_2, \ldots\ldots\ldots\ldots, A_n$ such that for given $\varepsilon > 0$

$$|u(x_1 + \Delta x_1, x_2 + \Delta x_2, \ldots\ldots, x_n + \Delta x_n) - u(x_1, x_2, \ldots., x_n)| = A_1\Delta x_1 + A_2\Delta x_2 + \ldots\ldots.. + A_n\Delta x_n + \varepsilon\rho$$

where $\rho = \sqrt{\sum_{i=1}^{n} \Delta x_i^2}$ & $\varepsilon \to 0$ whenever $\rho \to 0$.

Definition 4 (Partial derivative). If the increment ratio

$$\frac{u(x + \Delta x, y, z) - u(x, y, z)}{\Delta x}$$

tends to a unique limit as Δx tends to zero, this limit is called the **partial derivative** of u with respect to x and is written as $\frac{\partial u}{\partial x}$ or u_x.

Similarly, $\frac{\partial u}{\partial y}$ and $\frac{\partial u}{\partial z}$ can be defined.

So, if $u : R^n \to R$ be a function, we define a partial derivative as

$$\frac{\partial u}{\partial x_i} = \lim_{\Delta x_i \to 0} \frac{u(x_1, x_2, \ldots\ldots\ldots\ldots, x_i + \Delta x_i, \ldots\ldots.., x_n) - u(x_1, x_2, \ldots\ldots\ldots\ldots, x_n)}{\Delta x_i};\ i = 1, 2, \ldots\ldots., n..$$

The differential coefficients. If in the relation

$$\Delta u = A\Delta x + B\Delta y + C\Delta z + \in \rho$$

we suppose that $\Delta y = \Delta z = 0$, then, on the assumption that u is differentiable at the point (x, y, z),

$$\begin{aligned}\Delta u &= u(x + \Delta x, y, z) - u(x, y, z)\\ &= A\Delta x + \in \Delta x\end{aligned}$$

and by the taking limit as $\Delta x \to 0$, since $\in \to 0$ as $\Delta x \to 0$, we get $\frac{\partial u}{\partial x} = A$.

Similarly $\frac{\partial u}{\partial y} = B$ and $\frac{\partial u}{\partial z} = C$.

Hence, when the function u = u(x, y, z) is differentiable, the partial derivatives $\frac{\partial u}{\partial x}, \frac{\partial u}{\partial y}, \frac{\partial u}{\partial z}$ are respectively the differential coefficients A, B, C and so

$$\Delta u = \frac{\partial u}{\partial x}\Delta x + \frac{\partial u}{\partial y}\Delta y + \frac{\partial u}{\partial z}\Delta z + \in \rho$$

The differential of the dependent variable du is defined to be the principal part of Δu so that the above expression may be written as

$$\Delta u = du + \in \rho.$$

Now as in the case of functions of one variable, the differentials of the independent variables are identical with the arbitrary increment of these variables. If we write u = x, u = y, u = z respectively, it follows that

$$dx = \Delta x, dy = \Delta y, dz = \Delta z$$

Therefore, expression for du reduces to

$$du = \frac{\partial u}{\partial x}dx + \frac{\partial u}{\partial y}dy + \frac{\partial u}{\partial z}dz$$

Proposition 1. Let $f : R^n \to R$ be a function. If f is differentiable at a point $x = (x_1, x_2, \ldots\ldots, x_n) \in R^n$ then $f(x_1 + \Delta x_1, x_2 + \Delta x_2, \ldots\ldots, x_n + \Delta x_n) - f(x_1, x_2, \ldots\ldots, x_n)$

$$= \frac{\partial f}{\partial x_1}\Delta x_1 + \frac{\partial f}{\partial x_2}\Delta x_2 + \ldots\ldots + \frac{\partial f}{\partial x_n}\Delta x_n + \varepsilon\rho$$

where $\rho = \sqrt{\sum_{i=1}^{n} \Delta x_i^2}$ and $\varepsilon \to 0$ as $\rho \to 0$.

Proof. Since f is differentiable at a point $x = (x_1, x_2, \ldots\ldots, x_n)$, by definition of differentiability, there exists constants $A_1, A_2, \ldots\ldots, A_n$ such that, for given $\varepsilon > 0$

$$f(x_1 + \Delta x_1, x_2 + \Delta x_2, \ldots\ldots, x_n + \Delta x_n) - f(x_1, x_2, \ldots\ldots, x_n)$$
$$= A_1\Delta x_1 + A_2\Delta x_2 + \ldots\ldots + A_n\Delta x_n + \varepsilon\rho \qquad (*)$$

where $\rho = \sqrt{\sum \Delta x_i^2}$ and $\varepsilon \to 0$ as $\rho \to 0$.

Taking $\Delta x_j = 0$ for $j \neq i$ for some fixed $i = (1, 2, \ldots, n)$.

Thus, we have

$$\frac{f(x_1, x_2, \ldots\ldots, x_i + \Delta x_i, \ldots\ldots, x_n) - f(x_1, x_2, \ldots\ldots x_n)}{\Delta x_i} = A_i + \varepsilon$$

Taking $\Delta x_i \to 0$

$$\lim_{\Delta x_i \to 0} \frac{f(x_1, x_2, ..., x_i + \Delta x_i, ..., x_n) - f(x_1, x_2, ..., x_n)}{\Delta x_i} = A_i \quad \begin{bmatrix} \because \varepsilon \to 0 as \rho \to 0 \\ or \Delta x_i \to 0 \end{bmatrix}$$

$$\Rightarrow \frac{\partial f}{\partial x_i} = A_i \quad \text{(By definition of partial derivative)}$$

This is true for every $i = (1, 2, ..., n)$

$$\Rightarrow \frac{\partial f}{\partial x_1} = A_1, \frac{\partial f}{\partial x_2} = A_2, ..., \frac{\partial f}{\partial x_n} = A_n$$

Putting these value in equation (*), we get

$$f(x_1 + \Delta x_1, x_2 + \Delta x_2, ..., x_n + \Delta x_n) - f(x_1, x_2, ..., x_n)$$

$$= \frac{\partial f}{\partial x_1}\Delta x_1 + \frac{\partial f}{\partial x_2}\Delta x_2 + ... + \frac{\partial f}{\partial x_n}\Delta x_n + \varepsilon\rho$$

where $\rho = \left(\sum \Delta x_i\right)^{1/2}$ and $\varepsilon \to 0$ as $\rho \to 0$.

Remark 1. If the function $u = u(x_1, x_2, ..., x_n)$ is differentiable at point $(x_1, x_2, ..., x_n)$ then the partial derivative of u w.r.t. $x_1, x_2, ..., x_n$ certainly exist and are finite at this point, because by the above proposition, they are identical to constants $A_1, A_2, ..., A_n$ respectively.

However converse of this is not true, i.e., partial derivatives may exist at a point but the function need not be differential at that point.

In other words, we can say partial derivatives need not always be differential coefficients.

The distinction between derivatives and differential coefficients

We know that the necessary and sufficient condition that the function y = f(x) should be differentiable at the point x is that it possesses a finite definite derivative at that point. Thus for functions of one variable, the existence of derivative f'(x) implies the differentiability of f(x) at any given point.

For functions of more than one variable this is not true. If the function u = u(x, y, z) is differentiable at the point (x, y, z), the partial derivatives of u with respect to x, y and z certainly exist and are finite at this point, for then they are identical with differential coefficients A, B and C respectively. The partial derivatives, however, may exist at a point when the function is not differentiable at that point. In other words, the partial derivatives need not always be differential coefficients.

Example 1. Let f be a function defined by $f(x, y) = \frac{x^3 - y^3}{x^2 + y^2}$, where x and y are not simultaneously zero, f(0, 0) = 0.

If this function is differentiable at the origin, then, by definition,

$$f(h, k) - f(0, 0) = Ah + Bk + \in \eta \qquad (1)$$

where $\eta = \sqrt{h^2 + k^2}$ and $\in \to 0$ as $\eta \to 0$.

Putting $h=\eta\cos\theta, k=\eta\sin\theta$ in (1) and dividing through by η and taking limit as $\eta\to 0$, we get

$$\cos^3\theta-\sin^3\theta = A\cos\theta + B\sin\theta$$

which is impossible, since θ is arbitrary.

The function is therefore not differentiable at (0, 0). But the partial derivatives exist however, for

$$f_x(0,0)=\lim_{h\to 0}\frac{f(h,0)-f(0,0)}{h}=\lim_{h\to 0}\frac{h-0}{h}=1$$
$$f_y(0,0)=\lim_{k\to 0}\frac{f(0,k)-f(0,0)}{k}=\lim_{k\to 0}\frac{0-k}{k}=-1.$$

Example 2. Let $f(x,y)=\begin{cases}\dfrac{xy}{\sqrt{x^2+y^2}} & \text{if } x^2+y^2\neq 0\\ 0 & \text{if } x=0, y=0\end{cases}$.

Then $$f_x(0,0)=0=f_y(0,0)$$

and so partial derivatives exist. If it is different, then

$$df = f(h,k)-f(0,0) = Ah + Bk + \in\eta, \text{ where } A=f_x(0,0), B=f_y(0,0).$$

This yields

$$\frac{hk}{\sqrt{h^2+k^2}} = \in\sqrt{h^2+k^2}, \eta=\sqrt{h^2+k^2}$$

or $$hk = \in(h^2+k^2)$$

Putting k = mh, we get

$$mh^2 = \in h^2(1+m^2)$$

or $$\frac{m}{1+m^2}=\in$$

Hence $\lim_{k\to 0}\dfrac{m}{1+m^2}=0$, which is impossible. Hence the function is not differentiable at the origin.

Remark 2. (i)Thus the information given by the existence of the two first partial derivatives is limited. The values of $f_x(x,y)$ and $f_y(x,y)$ depend only on the values of f(x, y) along two lines through the point (x, y) respectively parallel to the axes of x and y. This information is incomplete and tells us nothing at all about the behavior of the function f(x, y) as the point (x, y) is approached along a line which is inclined to the axis of x at any given angle θ which is not equal to 0 or $\pi/2$.

(ii) Partial derivatives are also in general functions of x, y and z which may possess partial derivatives with respect to each of the three independent variables, we have the definition

a) $\dfrac{\partial}{\partial x}\left(\dfrac{\partial u}{\partial x}\right) = \lim_{\Delta x\to 0} \dfrac{u_x(x+\Delta x, y, z) - u_x(x, y, z)}{\Delta x}$

b) $\dfrac{\partial}{\partial y}\left(\dfrac{\partial u}{\partial x}\right) = \lim_{\Delta y\to 0} \dfrac{u_x(x, y+\Delta y, z) - u_x(x, y, z)}{\Delta y}$

c) $\dfrac{\partial}{\partial z}\left(\dfrac{\partial u}{\partial x}\right) = \lim_{\Delta z\to 0} \dfrac{u_x\left(x, y, z+\Delta z\right) - u_x\left(x, y, z\right)}{\Delta z}$

provided that each of these limits exist. We shall denote the second order partial derivatives by $\dfrac{\partial^2 u}{\partial x^2}$ or u_{xx}, $\dfrac{\partial^2 u}{\partial y \partial x}$ or u_{yx} and $\dfrac{\partial^2 u}{\partial z \partial x}$ or u_{zx}.

Similarly we may define higher order partial derivatives of $\dfrac{\partial u}{\partial y}$ and $\dfrac{\partial u}{\partial z}$.

The following example shows that certain second partial derivatives of a function may exist at a point at which the function is not continuous.

Example 3. Let $$\phi(x,\ y) = \begin{cases} \dfrac{x^3 + y^3}{x - y} & \text{when } (x, y) \neq (0, 0) \\ 0 & \text{when } (x, y) = (0, 0). \end{cases}$$

This function is discontinuous at the origin. To show this it is sufficient to prove that if the origin is approached along different paths, $\phi(x, y)$ does not tend to the same definite limit. For, if $\phi(x, y)$ were continuous at (0, 0), $\phi(x, y)$ would tend to zero (the value of the function at the origin) by whatever path the origin were approached.

Let the origin be approached along the three curves

$$\text{(i)}\ \ y = x - x^2 \qquad \text{(ii)}\ \ y = x - x^3 \qquad \text{(iii)}\ \ y = x - x^4;$$

Then we have

(i) $\phi(x,\ y) = \dfrac{2x^3 + 0(x^4)}{x^2} \to 0$ as $x \to 0$

(ii) $\phi(x,\ y) = \dfrac{2x^3 + 0(x^4)}{x^3} \to 2$ as $x \to 0$

(iii) $\phi(x,\ y) = \dfrac{2x^3 + 0(x^4)}{x^4} \to \infty$ as $x \to 0$

Certain partial derivatives, however, exist at (0, 0), for if ϕ_{xx} denote $\dfrac{\partial}{\partial x}\left(\dfrac{\partial \phi}{\partial x}\right)$ we have, for example

$$\phi_x(0,0) = \lim_{h\to 0}\frac{\phi(h,0)-\phi(0,0)}{h} = \lim_{h\to 0}\frac{h^2}{h} = 0$$

$$\phi_{xx}(0,0) = \lim_{h\to 0}\frac{\phi(h,0)-\phi_x(0,0)}{h} = \lim_{h\to 0}\frac{2h}{h} = 2,$$

since $\phi(x,0) = x^2, \phi_x(x,0) = 2x$ when $x \neq 0$.

The following example shows that u_{xy} is not always equal to u_{yx}.

Example 4. Let $$f(x,\ y) = \begin{cases} \dfrac{xy(x^2-y^2)}{x^2+y^2} & \text{when } (x,y)\neq(0,0) \\ 0 & \text{when } (x,y)=(0,0). \end{cases}$$

When the point (x, y) is not the origin, then

$$\frac{\partial f}{\partial x} = y\left[\frac{x^2-y^2}{x^2+y^2} + \frac{4x^2y^2}{(x^2+y^2)^2}\right] \tag{1}$$

$$\frac{\partial f}{\partial y} = x\left[\frac{x^2-y^2}{x^2+y^2} - \frac{4x^2y^2}{(x^2+y^2)^2}\right] \tag{2}$$

while at origin,

$$f_x(0,0) = \lim_{h\to 0}\frac{f(h,0)-f(0,0)}{h} = 0 \tag{3}$$

and similarly $f_y(0,0) = 0$.

From (1) and (2), we see that

$$f_x(0,y) = -y\ (y\neq 0) \quad \text{and} \quad f_y(x,0) = x\ (x\neq 0) \tag{4}$$

Now we have, using (3) and (4)

$$f_{xy}(0,0) = \lim_{h\to 0}\frac{f_y(h,0)-f_y(0,0)}{h} = \lim_{h\to 0}\frac{h}{h} = 1$$

$$f_{yx}(0,0) = \lim_{k\to 0}\frac{f_x(0,k)-f_x(0,0)}{k} = \lim_{h\to 0}\frac{-k}{k} = -1.$$

and so $f_{xy}(0,0) \neq f_{yx}(0,0)$.

Example 5. Prove that the function

$$f(x,y) = \sqrt{|xy|}$$

is not differentiable at the point (0, 0), but that $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$ both exist at the origin and have the value zero.

Hence deduce that these two partial derivatives are continuous except at the origin.

Solution. We have

$$\frac{\partial f}{\partial x}(0,0) = \lim_{h\to 0}\frac{f(h,0)-f(0,0)}{h} = 0$$
$$\frac{\partial f}{\partial y}(0,0) = \lim_{k\to 0}\frac{f(0,k)-f(0,0)}{k} = 0$$

If f(x, y) is differentiable at (0, 0), then we must have

$$f(h,k) = 0.h + 0.k + \in\sqrt{h^2+k^2}$$

where $\in\to 0$ as $\sqrt{h^2+k^2}\to 0$.

Now $\qquad \in = \frac{\sqrt{|hk|}}{\sqrt{h^2+k^2}}$

Putting $h=\rho\cos\theta, k=\rho\sin\theta,$ we get

$$\in = \sqrt{|\cos\theta\sin\theta|}$$

$\lim_{\rho\to 0}\in = \sqrt{|\cos\theta\sin\theta|} \Rightarrow \sqrt{|\cos\theta\sin\theta|} = 0$ which is impossible for arbitrary θ.

Hence, f is not differentiable.

Now, suppose that $(x,y)\neq(0,0).$ Then

$$\frac{\partial f}{\partial x} = \lim_{h\to 0}\frac{f(x+h,y)-f(x,y)}{h}$$
$$= \lim_{h\to 0}\frac{|(x+h)y|-|xy|}{h\left(\sqrt{|(x+h)y|}+\sqrt{|xy|}\right)} = \lim_{h\to 0}\sqrt{|y|}\frac{|x+h|-|x|}{h\left(\sqrt{|x+h|}+\sqrt{|x|}\right)}$$

Now, we can take h so small that x + h and x have the same sign. Hence the limit is $\frac{|y|}{2\sqrt{|xy|}}$ or $\frac{1}{2}\sqrt{\frac{|y|}{|x|}}$.

Similarly, $\frac{\partial f}{\partial y} = \frac{|x|}{2\sqrt{|xy|}}$ or $\frac{1}{2}\sqrt{\frac{|x|}{|y|}}$. Both of these are continuous except at (0,0). We now prove two theorems, the object of which is to set out precisely under what conditions it is allowable to assume that

$$f_{xy}(a,b) = f_{yx}(a,b).$$

Theorem 3 (Young). If (i) f_x and f_y exist in the neighbourhood of the point (a, b) and (ii) f_x and f_y are differentiable at (a, b); then

$$f_{xy} = f_{yx}.$$

Proof. We shall prove this theorem by taking equal increment h for both x and y and calculating $\Delta^2 f$ in two different ways, where

$$\Delta^2 f = f(a+h, b+h) - f(a+h, b) - f(a, b+h) + f(a, b).$$

Let

$$H(x) = f(x, b+h) - f(x, b)$$

Then

$$\Delta^2 f = H(a+h) - H(a).$$

Since f_x exists in the neighbourhood of (a, b), the function H(x) is derivable in $(a, a+h)$. Applying mean value theorem to H(x) for $0 < \theta < 1$, we obtain

$$H(a+h) - H(a) = hH'(a+\theta h)$$

Therefore

$$\begin{aligned} \Delta^2 f &= hH'(a+\theta h) \\ &= h[f_x(a+\theta h, b+h) - f_x(a+\theta h, b)] \end{aligned} \qquad (1)$$

By hypothesis (ii) of theorem, $f_x(x, y)$ is differentiable at (a, b) so that

$$f_x(a+\theta h, b+h) - f_x(a, b) = \theta h f_{xx}(a, b) + h f_{yx}(a, b) + \in' h$$

and

$$f_x(a+\theta h, b) - f_x(a, b) = \theta h f_{xx}(a, b) + \in'' h,$$

where $\in'$ and $\in''$ tend to zero as $h \to 0$. Thus, we get (on subtracting)

$$f_x(a+\theta h, b+h) - f_x(a+\theta h, b) = h f_{yx}(a, b) + (\in' - \in'')h$$

Putting this in (1), we obtain

$$\Delta^2 f = h^2 f_{yx} + \in_1 h^2 \qquad (2)$$

where $\in_1 = \in' - \in''$, so that $\in_1$ tends to zero with h.

Similarly, if we take

$$K(y) = f(a+h, y) - f(a, y)$$

Then we can show that

$$\Delta^2 f = h^2 f_{xy} + \in_2 h^2 \qquad (3)$$

where $\in_2 \to 0$ with h.

From (2) and (3), we have

$$\frac{\Delta^2 f}{h^2} = f_{yx}(a,b) + \in_1 = f_{xy}(a,b) + \in_2$$

Taking limit as $h \to 0$, we have

$$\lim_{h\to 0} \frac{\Delta^2 f}{h^2} = f_{yx}(a,b) = f_{xy}(a,b)$$

which establishes the theorem.

Theorem 4 (Schwarz). If (i) f_x, f_y, f_{yx} all exist in the neighbourhood of the point (a, b) and (ii) f_{yx} is continuous at (a, b); then f_{xy} also exist at (a, b) and $f_{xy} = f_{yx}$.

Proof. Let $(a+h, b+k)$ be point in neighbourhood of (a, b). Let (as in the above theorem)

$$\Delta^2 f = f(a+h, b+k) - f(a+h, b) - f(a, b+k) + f(a, b).$$

and

$$H(x) = f(x, b+k) - f(x, b)$$

so that we have

$$\Delta^2 f = H(a+h) - H(a).$$

Since f_x exists in the neighbourhood of (a, b), H(x) is derivable in $(a, a+h)$. Applying Mean value theorem to H(x) for $0 < \theta < 1$, we have

$$H(a+h) - H(a) = hH'(a+\theta h)$$

and therefore

$$\Delta^2 f = hH'(a+\theta h) = h[f_x(a+\theta h, b+k) - f_x(a+\theta h, b)].$$

Now, since f_{yx} exists in the neighbourhood of (a, b), the function f_x is derivable with respect to y in $(b, b+k)$. Applying mean value theorem, we have

$$\Delta^2 f = hkf_{yx}(a+\theta h, b+\theta' k), \quad 0 < \theta' < 1$$

That is

$$\frac{1}{h}\left[\frac{f(a+h, b+k) - f(a+h, b)}{k} - \frac{f(a, b+k) - f(a, b)}{k}\right] = f_{yx}(a+\theta h, b+\theta' k)$$

Taking limit as k tends to zero, we obtain

$$\frac{1}{h}[f_y(a+h, b) - f_y(a, b)] = \lim_{k\to 0} f_{yx}(a+\theta h, b+\theta' k) = f_{yx}(a+\theta h, b) \quad (1).$$

Since f_{yx} is given to be continuous at (a, b), we have

$$f_{yx}(a+\theta h, b) = f_{yx}(a,b) + \in,$$

where $\in \to 0$ and $h \to 0$.

Hence taking the limit $h \to 0$ in (1), we have

$$\lim_{h\to 0} \frac{f_y(a+h,b) - f_y(a,b)}{h} = \lim_{h\to 0}[f_{yx}(a,b) + \in]$$

that is, $$f_{xy}(a,b) = f_{yx}(a,b)$$

This completes the proof of the theorem.

Remark 3. The conditions of Young or Schwarz's Theorem are sufficient for $f_{xy} = f_{yx}$ but they are not necessary. For example, consider the function

$$f(x,\ y) = \begin{cases} \dfrac{x^2y^2}{x^2+y^2}, & (x,y) \neq (0,0) \\ 0 & , \quad (x,y) = (0,0). \end{cases}$$

We have

$$f_x(0,0) = \lim_{h\to 0} \frac{f(h,0) - f(0,0)}{h} = 0$$

$$f_y(0,0) = \lim_{k\to 0} \frac{f(0,k) - f(0,0)}{k} = 0$$

Also for $(x,y) \neq (0,0)$, we have

$$f_x(x,y) = \frac{(x^2+y^2)2xy^2 - x^2y^2.2x}{(x^2+y^2)^2} = \frac{2xy^4}{(x^2+y^2)^2}$$

$$f_y(x,y) = \frac{2x^4y}{(x^2+y^2)^2}$$

Again

$$f_{yx}(0,0) = \lim_{k\to 0} \frac{f_x(0,k) - f_x(0,0)}{k} = 0 \text{ and } f_{xy}(0,0) = 0$$

So that $f_{xy}(0,0) = f_{yx}(0,0)$.

For $(x,y) \neq (0,0)$, we have

$$f_{yx}(x,y) = \frac{8xy^3(x^2+y^2)^2 - 2xy^4 4y(x^2+y^2)}{(x^2+y^2)^4} = \frac{8x^3y^3}{(x^2+y^2)^3}$$

Putting y = mx, we can show that

$$\lim_{(x,y)\to(0,0)} f_{yx}(x,y) \neq 0 = f_{yx}(0,0)$$

so that f_{xy} is not continuous at (0, 0). Thus the condition of Schwarz's theorem is not satisfied.

To see that conditions of Young's theorem are also not satisfied, we notice that

$$f_{xx}(0,0) = \lim_{h \to 0} \frac{f_x(h,0) - f_x(0,0)}{h} = 0.$$

If f_x is differentiable at (0, 0), we should have

$$f_x(h,k) - f_x(0,0) = hf_{xx}(0,0) + kf_{yx}(0,0) + \in \eta$$

$$\frac{2hk^4}{(h^2+k^2)^2} = \in \eta,$$

where $\eta = \sqrt{h^2+k^2}$ and $\in \to 0$ as $\eta \to 0$.

Put $h = \eta \cos\theta, k = \eta \sin\theta$, then $\eta = \sqrt{h^2+k^2} = \rho$

so we have

$$\frac{2\rho\cos\theta.\rho^4\sin^4\theta}{\rho^4} = \in \rho$$

$$2\cos\theta.\sin^4\theta = \in$$

Taking limit as $\rho \to 0$, we have

$$2\cos\theta.\sin^4\theta = 0$$

which is impossible for arbitrary θ.

3.4 References

1. Walter Rudin, Principles of Mathematical Analysis (3rd edition) McGraw-Hill, Kogakusha, 1976, International Student Edition.
2. T. M. Apostol, Mathematical Analysis, Narosa Publishing House, New Delhi, 1974.
3. H.L. Royden, Real Analysis, Macmillan Pub. Co., Inc. 4th Edition, New York, 1993.
4. G. De Barra, Measure Theory and Integration, Wiley Eastern Limited, 1981.
5. R.R. Goldberg, Methods of Real Analysis, Oxford & IBH Pub. Co. Pvt. Ltd, 1976.
6. R. G. Bartle, The Elements of Real Analysis, Wiley International Edition, 2011.
7. S.C. Malik and Savita Arora, Mathematical Analysis, New Age International Limited, New Delhi, 2012.

UNIT – IV

TAYLOR THEOREM

Structure

4.0 Introduction

In this unit, we study most important mathematical tool of analysis i.e. Taylor theorem. As we know, Taylor series is an expression of a function as an infinite series whose terms are expressed in term of the values of the function's derivatives at a single point. Also we shall be mainly concerned with the applications of differential calculus to functions of more than one variable such as how to find stationary points and extreme values of implicit functions, implicit function theorem, Jacobian and its properties etc.

4.1 Unit Objectives

After going through this unit, one will be able to

- solve Taylor series expansions.
- find the stationary points and extreme values of implicit functions.
- understand Jacobian and its properties.
- know about the local character of Implicit function i.e. the implicit function is a unique solution of a function $f(x, y)=0$ in a certain neighbourhood.

4.2 Taylor Theorem

In view of Taylor's theorem for functions of one variable, it is not unnatural to expect the possibility of expanding a function of more than one variable $f(x+h, y+k, z+m)$, in a series of ascending powers of h, k, m. To fix the ideas, consider a function of two variables only; the reasoning in general case is precisely the same.

Theorem 1 (Taylor's theorem). If f(x, y) and all its partial derivatives of order n are finite and continuous for all point (x, y) in domain $a \le x \le a+h, b \le y \le b+k,$ then

$$f(a+h, b+k) = f(a,b) + df(a,b) + \frac{1}{2!}d^2 f(a,b) + \ldots + \frac{1}{(n-1)!}d^{n-1}f(a,b) + R_n$$

where $$R_n = \frac{1}{n!}d^n f(a+\theta h, b+\theta k), 0 < \theta < 1.$$

Proof. Consider a circular domain of centre (a, b) and radius large enough for the point $(a+h, b+k)$ to be also with in domain. Suppose that f(x, y) is a function such that all the partial derivatives of order n of f(x, y) are continuous in the domain. Write

$$x = a + ht, y = b + kt,$$

so that, as t ranges from 0 to 1, the point (x, y) moves along the line joining the point (a, b) to the point $(a+h, b+k)$; then

$$f(x, y) = f(a+ht, b+kt) = \phi(t).$$

Now, $\phi'(t) = \frac{\partial f}{\partial x}.\frac{dx}{dt} + \frac{\partial f}{\partial y}.\frac{dy}{dt} = h\frac{\partial f}{\partial x} + k\frac{\partial f}{\partial y} = df$

and

$$\begin{aligned}
\phi''(t) &= \frac{\partial}{\partial x}\left(h\frac{\partial f}{\partial x} + k\frac{\partial f}{\partial y}\right)\frac{dx}{dt} + \frac{\partial}{\partial y}\left(h\frac{\partial f}{\partial x} + k\frac{\partial f}{\partial y}\right)\frac{dy}{dt} \\
&= h\frac{\partial^2 f}{\partial x^2}\frac{dx}{dt} + k\frac{\partial^2 f}{\partial x \partial y}\frac{dx}{dt} + h\frac{\partial^2 f}{\partial y \partial x}\frac{dy}{dt} + k\frac{\partial^2 f}{\partial y^2}\frac{dy}{dt} \\
&= \left(h^2\frac{\partial^2 f}{\partial x^2} + hk\frac{\partial^2 f}{\partial x \partial y} + hk\frac{\partial^2 f}{\partial y \partial x} + k^2\frac{\partial^2 f}{\partial y^2}\right) \\
&= \left(h^2\frac{\partial^2}{\partial x^2} + 2hk\frac{\partial^2}{\partial x \partial y} + k^2\frac{\partial^2}{\partial y^2}\right)f \qquad (by\ Schwarz's\ theorem) \\
&= \left(h\frac{\partial}{\partial x} + k\frac{\partial}{\partial y}\right)^2 f(a+ht, b+kt)
\end{aligned}$$

and hence, similarly we get

$$\phi''(t) = d^2 f, \ldots, \phi^{(n)}(t) = d^n f$$

Also, $\phi(t)$ and its n derivatives are continuous functions of t in the interval $0 \le t \le 1$, and so, by Maclaurin's theorem

$$\phi(t) = \phi(0) + t\phi'(0) + \frac{t^2}{2!}\phi''(0) + \ldots + \frac{t^n}{n!}\phi^{(n)}(\theta t) \qquad (1)$$

where $0 < \theta < 1$. Now put t = 1 and observe that

$$\begin{aligned} &\phi(1) = f(a+h, b+k), \\ &\phi(0) = f(a,b), \\ &\phi'(0) = df(a,b), \\ &\phi''(0) = d^2 f(a,b), \\ &\ldots\ldots \\ &\phi^{(n)}(\theta t) = d^n f(a+\theta h, b+\theta k). \end{aligned}$$

It follows immediately from (1) that

$$f(a+h, b+k) = f(a,b) + df(a,b) + \frac{1}{2!}d^2 f(a,b) + \ldots + \frac{1}{(n-1)!}d^{n-1}f(a,b) + R_n \qquad (2)$$

where $$R_n = \frac{1}{n!}d^n f(a+\theta h, b+\theta k), 0 < \theta < 1.$$

Here, we assumed that all the partial derivatives of order n are continuous in the domain. Taylor expansion does not necessarily hold if these derivatives are not continuous.

Remark 1. If we put $a = b = 0, h = x, k = y,$ from the equation (2), we get

$$f(x,y) = f(0,0) + df(0,0) + \frac{1}{2!}d^2 f(0,0) + \ldots + \frac{1}{(n-1)!}d^{n-1}f(0,0) + R_n$$

where $$R_n = \frac{1}{n!}d^n f(\theta x, \theta y), 0 < \theta < 1.$$

This is known as Maclaurin's theorem.

2. If we put $a + h = x, \quad b + k = y,$ we get

$$f(x,y) = f(a,b) + \left[(x-a)\frac{\partial}{\partial x} + (y-b)\frac{\partial}{\partial y}\right] f(a,b) + \ldots\ldots\ldots$$

$$+ \frac{1}{(n-1)!}\left[(x-a)\frac{\partial}{\partial x} + (y-b)\frac{\partial}{\partial y}\right]^{n-1} f(a,b) + R_n,$$

where $$R_n = \frac{1}{n!}\left[(x-a)\frac{\partial}{\partial x} + (y-b)\frac{\partial}{\partial y}\right]^{n} f(a+(x-a)\theta, b+(y-b)\theta).$$

This is called Taylor's expansion of $f(x,y)$ about the point (a,b) in power of $(x-a)$ and $(y-b)$.

Example 1. If $f(x, y)=\sqrt{|xy|}$, prove that Taylor's expansion about the point (x, x) is not valid in any domain which includes the origin.

Solution. Given that $f(x,y)=\sqrt{|xy|}$.

We find $$f_x(0,0)=\lim_{h\to 0}\frac{f(h,0)-f(0,0)}{h}=0$$

$$f_y(0,0)=\lim_{k\to 0}\frac{f(0,k)-f(0,0)}{k}=0$$

Now, $$f_x(x,y)=\begin{cases}\frac{1}{2}\sqrt{\frac{|y|}{|x|}}, x>0\\ -\frac{1}{2}\sqrt{\frac{|y|}{|x|}}, x<0\end{cases}$$

Also $$f_y(x,y)=\begin{cases}\frac{1}{2}\sqrt{\frac{|x|}{|y|}}, y>0\\ -\frac{1}{2}\sqrt{\frac{|x|}{|y|}}, y<0\end{cases}$$

Thus, $$f_x(x,x)=f_y(x,x)=\begin{cases}\frac{1}{2}, x>0\\ -\frac{1}{2}, x<0\end{cases}$$

Now, Taylor's expansion about (x,x) for $n=1$ is

$$f(x+h,x+h)=f(x,x)+h\left\{f_x(x+\theta h,x+\theta h)+f_y(x+\theta h,x+\theta h)\right\}$$

$$|x+h|=\begin{cases}|x|+h, x+\theta h>0\\ |x|-h, x+\theta h<0\\ |x|, x+\theta h=0.\end{cases} \tag{1}$$

If the domain $((x,x),(x+h,x+h))$ contains origin then x and $x+h$ must be of opposite sign i.e.

$|x+h|=x+h$, $\quad |x|=-x$

or $\quad |x+h|=-(x+h)$, $\quad |x|=x$

under these conditions none of the equality in (1) holds.

Hence the expansion is not possible because partial derivatives f_x and f_y are not continuous in any domain which contains origin.

($\because$ Partial derivatives f_x, f_y are not continuous at origin and therefore Taylor's theorem is not necessary valid).

Example 2. Expand $x^2y+3y-2$ in power of $(x-1),(y+2)$.

Solution. Let us use Taylor's expansion with $a=1, b=-2$.

Then, $f(x,y)=x^2y+3y-2$, $f(1,-2)=-10$

$f_x(x,y)=2xy$, $f_x(1,-2)=-4$

$f_y(x,y)=x^2+3$, $f_y(1,-2)=4$

$f_{xx}(x,y)=2y$, $f_{xx}(1,-2)=-4$

$f_{xy}(x,y)=2x$, $f_{xy}(1,-2)=2$

$f_{yy}(x,y)=0$, $f_{yy}(1,-2)=0$

$f_{xxx}(x,y)=0$, $f_{xxx}(1,-2)=0$

$f_{yyy}(x,y)=0$, $f_{yyy}(1,-2)=0$

$f_{yxx}(x,y)=2$, $f_{yxx}(1,-2)=2$

$f_{xxy}(x,y)=2$, $f_{xxy}(1,-2)=2$.

All higher derivatives are zero. Thus, we have

$$x^2y+3y-2=-10-4(x-1)+4(y+2)-2(x-1)^2+2(x-1)(y+2)+(x-1)^2(y+2).$$

4.3 Explicit and Implicit Functions

The explicit function is one which is given in the independent variable. On the other hand, implicit functions are usually given in terms of both dependent and independent variables. Here we read in details:

Explicit function

If we consider set of n independent variables $x_1, x_2, x_3......, x_n$ and one dependent variable u, the equation

$$u=f\left(x_1,x_2,x_3......,x_n\right) \qquad (*)$$

denotes the functional relation. In this case if $y_1, y_2, y_3......, y_n$ are the n arbitrarily assigned values of the independent variables, the corresponding values of the dependent variable u are determined by the functional relation.

The function represented by equation $(*)$ is an **Explicit function** but where several variables are involved, then it is difficult to express one variable explicitly in terms of the others. Thus most of the functions of more than one variable are implicit function, that is to say we are given a functional relation

$$\phi\left(x_1, x_2, x_3......, x_n\right) = 0$$

connecting the n variables $x_1, x_2, x_3......, x_n$ and is not in general possible to solve this equation to find an explicit function which expresses one of these variables say x_1, in terms of the other $n-1$ variables.

Implicit function

Let $$\mathrm{F}(\mathrm{x}_1, \mathrm{x}_2, ..., \mathrm{x}_n, \mathrm{u}) = 0 \qquad (1)$$

be a functional relation between the $\mathrm{n}+1$ variables $\mathrm{x}_1, \mathrm{x}_2, ..., \mathrm{x}_n, \mathrm{u}$ and let $\mathrm{x}_1 = \mathrm{a}_1, \mathrm{x}_2 = \mathrm{a}_2, ...,\ \mathrm{x}_n = \mathrm{a}_n$ be a set of values such that the equation.

$$\mathrm{F}(\mathrm{a}_1, \mathrm{a}_2, ..., \mathrm{a}_n, \mathrm{u}) = 0 \qquad (2)$$

is satisfied for at least one value of u, that is equation (2) in u has at least one root. We may consider u as a function of the x's : $\mathrm{u} = \phi(\mathrm{x}_1, \mathrm{x}_2, ..., \mathrm{x}_n)$ defined in a certain domain, where $\phi(\mathrm{x}_1, \mathrm{x}_2, ..., \mathrm{x}_n)$ has assigned to it at any point $(\mathrm{x}_1, \mathrm{x}_2, ..., \mathrm{x}_n)$ the roots u of the equation (1) at this point. We say that u is the implicit function defined by (1). It is, in general, a many valued function.

More generally, consider the set of equations

$$\mathrm{F}_\mathrm{p}(\mathrm{x}_1, \mathrm{x}_2, ..., \mathrm{x}_n, \mathrm{u}_1, ..., \mathrm{u}_m) = 0 \ (\mathrm{p} = 1, 2, ..., \mathrm{m}) \qquad (3)$$

between the $\mathrm{n}+\mathrm{m}$ variables $\mathrm{x}_1, ..., \mathrm{x}_n, \mathrm{u}_1, ..., \mathrm{u}_m$ and suppose that the set of equations (3) are such that there are points $(\mathrm{x}_1, \mathrm{x}_2, ..., \mathrm{x}_n)$ for which these m equations are satisfied for at least one set of values $\mathrm{u}_1, \mathrm{u}_2, ..., \mathrm{u}_m$ We may consider the u's as function of x's.

$$\mathrm{u}_\mathrm{p} = \phi_\mathrm{p}(\mathrm{x}_1, \mathrm{x}_2, ..., \mathrm{x}_n) \ (\mathrm{p} = 1, 2, ..., \mathrm{m})$$

where the function ϕ have assigned to them at the point $(\mathrm{x}_1, \mathrm{x}_2, ..., \mathrm{x}_n)$ the values of the roots $\mathrm{u}_1, \mathrm{u}_2, ..., \mathrm{u}_m$ at this point. We say that $\mathrm{u}_1, \mathrm{u}_2, ..., \mathrm{u}_m$ constitute a system of implicit functions defined by the set of equation (3). These functions are in general many valued.

Definition 1 (Implicit function of two variables). Let $f(x, y)$ be a function of two variables and $y = \phi(x)$ be a function of x such that for every value of x for which $\phi(x)$ is defined, $f(x, \phi(x))$ vanishes identically i.e., $y = \phi(x)$ is a root of the functional equation $f(x, y) = 0$. Then, $y = \phi(x)$ is an implicit function defined by the functional equation $f(x, y) = 0$.

4.3.1 Implicit function theorem.

This theorem tells us that whenever we can solve the approximating linear equation for y as a function of x, then the original equation defines y implicitly as a function of x. This theorem also known as Existence theorem.

Theorem 1 (Implicit function theorem). Let F(u, x, y) be a continuous function of variables u, x, y. Suppose that

(i) $F(u_0, a, b) = 0$;

(ii) F(u, a, b) is differentiable at (u_0, a, b);

(iii) The partial derivative $\frac{\partial F}{\partial u}(u_0, a, b) \neq 0$.

Then there exists at least one function u = u(x, y) reducing to u_0 at the point (a, b) and which, in the neighbourhood of this point, satisfies the equation F(u, x, y) = 0 identically.

Also, every function u which possesses these two properties is continuous and differentiable at the point (a, b).

Proof. Since $F(u_0, a, b) = 0$ and $\frac{\partial F}{\partial u}(u_0, a, b) \neq 0$, the function F is either an increasing or decreasing function of u when $u = u_0$. Thus there exists a positive number δ such that $F(u_0 - \delta, a, b)$ and $F(u_0 + \delta, a, b)$ have opposite signs. Since F is given to be continuous, a positive number η can be found so that the functions

$$F(u_0 - \delta, x, y) \text{ and } F(u_0 + \delta, x, y)$$

the values of which may be as near as we please to

$$F(u_0 - \delta, a, b) \text{ and } F(u_0 + \delta, a, b)$$

will also have opposite signs so long as $|x - a| < \eta$ and $|y - b| < \eta$.

Let x, y be any two values satisfying the above conditions. Then F(u, x, y) is a continuous function of u which changes sign between $u_0 - \delta$ and $u_0 + \delta$ and so vanishes somewhere in this interval. Thus for these x and y there is a u in $[u_0 - \delta, u_0 + \delta]$ for which F(u, x, y) = 0. Thus u is a function of x and y, say u(x, y) which reduces to u_0 at the point (a, b).

Suppose that $\Delta u, \Delta x, \Delta y$ are the increments of such function u and of the variables x and y measured from the point (a, b). Since F is differentiable at (u_0, a, b) we have

$$\Delta F = [F_u(u_0, a, b) + \in]\Delta u + [F_x(u_0, a, b) + \in']\Delta x + [F_y(u_0, a, b) + \in'']\Delta y = 0.$$

Since $\Delta F = 0$ because of F = 0. The numbers $\in, \in', \in''$ tend to zero with $\Delta u, \Delta x$ & Δy and can be made as small as we please with δ & η. Let δ and η be so small that the numbers $\in, \in', \in''$ are all less than $\frac{1}{2}\left|F_u(u_0, a, b)\right|$, which is not zero by our hypothesis. The above equation then shows that $\Delta u \to 0$ as $\Delta x \to 0$ and $\Delta y \to 0$ which means that the function $u = u(x, y)$ is continuous at (a, b).

Moreover, we have

$$\Delta u = -\frac{\left[F_x\left(u_0,a,b\right)+\in'\right]\Delta x+\left[F_y\left(u_0,a,b\right)+\in''\right]\Delta y}{F_u\left(u_0,a,b\right)+\in}$$

$$= -\frac{F_x\left(u_0,a,b\right)}{F_u\left(u_0,a,b\right)}\Delta x-\frac{F_y\left(u_0,a,b\right)}{F_u\left(u_0,a,b\right)}\Delta y+\in_1\ \Delta x+\in_2\ \Delta y,$$

$\in_1$ and $\in_2$ tending to zero as Δx and Δy tend to zero.

Hence u is differentiable at (a, b).

Corollary 1. If $\frac{\partial F}{\partial u}$ exists and is not zero in the neighbourhood of the point $\left(u_0,a,b\right)$, the solution u of the equation F = 0 is unique. Suppose that there are two solutions u_1 and u_2. Then we should have, by mean value theorem, for $u_1 < u' < u_2$

$$0 = F\left(u_1,x,y\right)-F\left(u_2,x,y\right)=\left(u_1-u_2\right)F_u\left(u',x,y\right),$$

and so $F_u\left(u,x,y\right)$ would vanish at some point in the neighbourhood of $\left(u_0,a,b\right)$ which is contrary to our hypothesis.

Corollary 2. If $F(u,x,y)$ is differentiable in the neighbourhood of $\left(u_0,a,b\right)$, the function $u = u(x,y)$ is differentiable in the neighbourhood of the point (a, b).

This is immediate, because the preceding proof is then application at every point (u,x,y) in that neighbourhood.

4.3.2 Inverse function theorem.

Corollary 1 is of great importance, for a function of two variables only, $F(u,x)=0$ and taking $F(u,x)=f(u)-x$, we can express the fundamental theorem on inverse functions as follows:

Theorem 1 (Inverse function theorem). If, in the neighbourhood of $u=u_0$, the function $f(u)$ is a continuous function of u and if

(i) $f\left(u_0\right)=a$

(ii) $f'(u)\neq 0$

in the neighbourhood of the point $u=u_0$, then there exists a unique continuous function $u=\phi(x)$, which is equal to u_0 when x = a, and which satisfies identically the equation

$$f(u)-x=0,$$

in the neighbourhood at the point x = a.

The function $u=\phi(x)$ thus defined is called the inverse function of $x=f(u)$.

4.4 Higher Order Differentials

Let $z = f(x, y)$ be a function of two independent variables x and y defined in a certain domain and let it be differentiable at the point (x, y) of the domain. The first differential coefficient of z at the point (x, y) is defined as

$$dz = \frac{\partial z}{\partial x} dx + \frac{\partial z}{\partial y} dy \qquad (1)$$

and if $\frac{\partial z}{\partial x}$ and $\frac{\partial z}{\partial y}$ are differentiable at the point (x, y), then the differential coefficient of dz is called second differential coefficient of z and is denoted by $d^2 z$ and is given by

$$d^2 z = d\left(\frac{\partial z}{\partial x} dx + \frac{\partial z}{\partial y} dy\right)$$

$$= d\left(\frac{\partial z}{\partial x}\right) dx + d\left(\frac{\partial z}{\partial y}\right) dy \qquad (2)$$

Now, $$d\left(\frac{\partial z}{\partial x}\right) = \frac{\partial^2 z}{\partial x^2} dx + \frac{\partial^2 z}{\partial y \partial x} dy$$

and $$d\left(\frac{\partial z}{\partial y}\right) = \frac{\partial^2 z}{\partial x \partial y} dx + \frac{\partial^2 z}{\partial y^2} dy.$$

Putting these values in (2), we get

$$d^2 z = \frac{\partial^2 z}{\partial x^2} (dx)^2 + 2\frac{\partial^2 z}{\partial y \partial x} dxdy + \frac{\partial^2 z}{\partial y^2} (dy)^2$$

Thus, $$d^2 z = \left(\frac{\partial}{\partial x} dx + \frac{\partial}{\partial y} dy\right)^2 z$$

Similarly, $$d^3 z = \left(\frac{\partial}{\partial x} dx + \frac{\partial}{\partial y} dy\right)^3 z$$

Proceeding in this manner, we define the successive differential coefficients $d^4 z, d^5 z, \ldots\ldots\ldots.$

Thus, the differential coefficient of nth order $d^n z$ exists if $d^{n-1} z$ is differentiable i.e. if all the partial derivatives of (n-1)th order are differentiable. Thus, by mathematical induction, we have

$$d^n z = \frac{\partial^n z}{\partial x^n} (dx)^n + n\frac{\partial^n z}{\partial x^{n-1} \partial y} dx^{n-1} dy + \frac{n(n-1)}{2!} \frac{\partial^n z}{\partial x^{n-2} \partial y^2} dx^{n-2} (dy)^2 + \ldots\ldots\ldots + \frac{\partial^n z}{\partial y^n} (dy)^n$$

$$=\left(\frac{\partial}{\partial x}dx+\frac{\partial}{\partial y}dy\right)^n z\,.$$

4.4.1. Choice of independent variables

$$\text{Let } F(x,y,z)=0 \qquad (1)$$

Differentiate (1), we get

$$\frac{\partial F}{\partial x}dx+\frac{\partial F}{\partial y}dy+\frac{\partial F}{\partial z}dz=0 \qquad (2)$$

Now, if z is dependent on the two independent variables x and y in such a way that the equation $F(x,y,z)=0$ is satisfied by $z=z(x,y)$, then

$$dz=\frac{\partial z}{\partial x}dx+\frac{\partial z}{\partial y}dy \qquad (3)$$

Now, equation (2) can be written as

$$dz=-\frac{F_x}{F_z}dx-\frac{F_y}{F_z}dy \qquad (4)$$

Comparing (3) and (4), we get

$$\frac{\partial z}{\partial x}=-\frac{F_x}{F_z}, \qquad \frac{\partial z}{\partial y}=-\frac{F_y}{F_z}$$

Similarly, if x is dependent on y and z then

$$\frac{\partial x}{\partial y}=-\frac{F_y}{F_x}, \qquad \frac{\partial x}{\partial z}=-\frac{F_z}{F_x}$$

Similarly, if y is dependent on z and x, then

$$\frac{\partial y}{\partial x}=-\frac{F_x}{F_y}, \qquad \frac{\partial y}{\partial z}=-\frac{F_z}{F_y}.$$

4.4.2 Higher order derivatives of implicit functions

Let $f(x,y,z)=0$ be a functional relation where z is dependent variable such that $z=z(x,y)$.

We denote the partial derivatives $\frac{\partial z}{\partial x},\frac{\partial z}{\partial y},\frac{\partial^2 z}{\partial x^2},\frac{\partial^2 z}{\partial x\partial y},\frac{\partial^2 z}{\partial y^2}$ by p, q, r, s, t respectively.

Now, we suppose that x is dependent variable so that $x=x(y,z)$. Then, we will show that how to express partial derivatives of first and second order w.r.t. y and z in terms of p, q, r, s and t.

Since $\qquad z=z(x,y)$

$$\Rightarrow \qquad dz = \frac{\partial z}{\partial x}dx + \frac{\partial z}{\partial y}dy. \qquad (1)$$

Now, we differentiate (1), taking x as dependent variable, dy and dz as constant so that

$$\Rightarrow \qquad 0 = d\left(\frac{\partial z}{\partial x}\right)dx + \frac{\partial z}{\partial x}d^2x + d\left(\frac{\partial z}{\partial y}\right)dy$$

$$= \frac{\partial^2 z}{\partial x^2}(dx)^2 + \frac{\partial^2 z}{\partial x \partial y}dydx + \frac{\partial z}{\partial x}d^2x + \frac{\partial^2 z}{\partial x \partial y}dxdy + \frac{\partial^2 z}{\partial y^2}(dy)^2$$

$$= r(dx)^2 + 2sdx.dy + t(dy)^2 + pd^2x. \qquad (2)$$

Now, from (1)

$$dz = pdx + qdy \qquad (3)$$

$$\Rightarrow \qquad dx = \frac{1}{p}(dz - qdy) \qquad (4)$$

Now, putting the value of dx in (2), we get

$$0 = r\left[\frac{1}{p}(dz - qdy)\right]^2 + 2s\left[\frac{1}{p}(dz - qdy)\right]dy + t(dy)^2 + pd^2x$$

$$\Rightarrow \qquad -pd^2x = r.\frac{1}{p^2}\left[(dz)^2 + q^2(dy)^2 - 2qdz.dy\right] + 2s\left[\frac{1}{p}\left(dzdy - q\left(dy\right)^2\right)\right] + t(dy)^2$$

$$= \frac{r}{p^2}(dz)^2 + \left(\frac{rq^2}{p^2} - \frac{2sq}{p} + t\right)(dy)^2 + \left(\frac{2s}{p} - \frac{2qr}{p^2}\right)dzdy$$

$$= \frac{r}{p^2}(dz)^2 + \frac{\left(rq^2 - 2spq + tp^2\right)}{p^2}(dy)^2 + \frac{(2sp - 2qr)}{p^2}dzdy$$

$$\Rightarrow \qquad d^2x == -\frac{r}{p^3}(dz)^2 + \frac{\left(2pqs - rq^2 - tp^2\right)}{p^3}(dy)^2 + \frac{(2qr - 2sp)}{p^3}dzdy. \qquad (5)$$

From (4), we have

$$\frac{\partial x}{\partial z} = \text{Coefficient of } dz \text{ in } (4) = + \frac{1}{p};$$

$$\frac{\partial x}{\partial y} = \text{Coefficient of } dy \text{ in } (4) = -\frac{q}{p}.$$

From (5), we have

$$\frac{\partial^2 x}{\partial z^2} = \text{Coefficient of } (dz)^2 \text{ in (5)} = -\frac{r}{p^3}$$

$$\frac{\partial^2 x}{\partial y^2} = \text{Coefficient of } (dy)^2 \text{ in (5)} = \frac{2pqs - rq^2 - p^2 t}{p^3}$$

$$\frac{\partial^2 x}{\partial y \partial z} = \frac{1}{2} \text{Coefficient of } dydz \text{ in (5)} = \frac{1}{2}\frac{(2qr - 2sp)}{p^3}$$

$$= \frac{qr - sp}{p^3}.$$

4.5 Change of Variables

In problems involving change of variables it is frequently required to transform a particular expression involving a combination of derivatives with respect to a set of variables, in term of derivatives with respect to another set of variables.

Example1. Let w be a function of two variables x and y, then transform the expression $\frac{\partial^2 w}{\partial x^2} + \frac{\partial^2 w}{\partial y^2}$ by the formula of polar transformation $x = u\cos v$, $y = u\sin v$.

Solution. Here, $x = x(u,v)$

$$dx = \frac{\partial x}{\partial u}du + \frac{\partial x}{\partial v}dv$$

$$= \cos v.du - u\sin v.dv \qquad (1)$$

Since $y = y(u,v)$

$$dy = \frac{\partial y}{\partial u}du + \frac{\partial y}{\partial v}dv$$

$$= \sin v.du + u\cos v.dv \qquad (2)$$

Multiplying (1) by $\cos v$ and (2) by $\sin v$ and adding, we get

$$du = \cos v(dx) + \sin v(dy) \qquad (3)$$

Multiplying (1) by $\sin v$ and (2) by $\cos v$ and subtracting, we get

$$dv = -\frac{\sin v}{u}(dx) + \frac{\cos v}{u}(dy) \qquad (4)$$

From (3) and (4), we get

$$\frac{\partial u}{\partial x} = \cos v, \qquad \frac{\partial u}{\partial y} = \sin v$$

$$\frac{\partial v}{\partial x}=-\frac{\sin v}{u}, \quad \frac{\partial v}{\partial y}=\frac{\cos v}{u}$$

Now, $$\frac{\partial w}{\partial x}=\frac{\partial w}{\partial u}.\frac{\partial u}{\partial x}+\frac{\partial w}{\partial v}.\frac{\partial v}{\partial x}$$

$$=\cos v\frac{\partial w}{\partial u}+\left(-\frac{\sin v}{u}\right)\frac{\partial w}{\partial v}$$

$$=\left(\cos v\frac{\partial}{\partial u}-\frac{\sin v}{u}\frac{\partial}{\partial v}\right)w \qquad (5)$$

Similarly $$\frac{\partial w}{\partial y}=\left(\sin v\frac{\partial}{\partial u}+\frac{\cos v}{u}\frac{\partial}{\partial v}\right)w \qquad (6)$$

Now $$\frac{\partial^2 w}{\partial x^2}=\frac{\partial}{\partial x}\left(\frac{\partial w}{\partial x}\right)$$

$$=\left(\cos v\frac{\partial}{\partial u}-\frac{\sin v}{u}\frac{\partial}{\partial v}\right)\left(\cos v\frac{\partial w}{\partial u}-\frac{\sin v}{u}\frac{\partial w}{\partial v}\right)$$

$$=\cos^2 v\frac{\partial^2 w}{\partial u^2}-\frac{\sin v\cos v}{u}\frac{\partial^2 w}{\partial u\partial v}+\frac{\sin v\cos v}{u^2}\frac{\partial w}{\partial v}-\frac{\sin v\cos v}{u}\frac{\partial^2 w}{\partial u\partial v}$$

$$+\frac{\sin^2 v}{u}\frac{\partial w}{\partial u}+\frac{\sin v\cos v}{u^2}\frac{\partial w}{\partial v}+\frac{\sin^2 v}{u^2}\frac{\partial^2 w}{\partial v^2} \qquad (7)$$

$$=\cos^2 v\frac{\partial^2 w}{\partial u^2}-\frac{2\sin v\cos v}{u}\frac{\partial^2 w}{\partial u\partial v}+\frac{\sin^2 v}{u^2}\frac{\partial^2 w}{\partial u\partial v}+\frac{\sin^2 v}{u}\frac{\partial w}{\partial u}+\frac{2\sin v\cos}{u^2}\frac{\partial w}{\partial u}$$

Similarly $$\frac{\partial^2 w}{\partial y^2}=\sin^2 v\frac{\partial^2 w}{\partial u^2}+\frac{\sin v\cos v}{u}\frac{\partial^2 w}{\partial u\partial v}-\frac{\sin v\cos v}{u^2}\frac{\partial w}{\partial v}+\frac{\cos^2 v}{u}\frac{\partial w}{\partial u}$$

$$+\frac{\sin v\cos v}{u}\frac{\partial^2 w}{\partial u\partial v}-\frac{\cos v\sin v}{u^2}\frac{\partial w}{\partial v}+\frac{\cos^2 v}{u^2}\frac{\partial^2 w}{\partial v^2} \qquad (8)$$

Adding (7) and (8), we get

$$\frac{\partial^2 w}{\partial x^2}+\frac{\partial^2 w}{\partial y^2}=\frac{\partial^2 w}{\partial u^2}+\frac{1}{u}\frac{\partial w}{\partial u}+\frac{1}{u^2}\frac{\partial^2 w}{\partial v^2}.$$

Example 2. Transform the expression

$$\left(x\frac{\partial z}{\partial x}+\frac{\partial z}{\partial y}\right)^2+\left(a^2-x^2-y^2\right)\left\{\left(\frac{\partial z}{\partial x}\right)^2+\left(\frac{\partial z}{\partial y}\right)^2\right\}$$

by the substitution $x=r\cos\theta, y=r\sin\theta$.

Solution. We wish to express z as a function of x and y where x and y are the functions of r and θ i.e. $z=z(x,y)$ and given $x=x(r,\theta)$, $y=y(r,\theta)$.

$$dx = \frac{\partial x}{\partial r}dr + \frac{\partial x}{\partial \theta}d\theta$$

$$= \cos\theta dr + (-\sin\theta.r)d\theta$$

$$= \cos\theta dr - \sin\theta.rd\theta \qquad (1)$$

Similarly $$dy = \frac{\partial y}{\partial r}dr + \frac{\partial y}{\partial \theta}d\theta$$

$$= \sin\theta dr + r\cos\theta d\theta \qquad (2)$$

Multiplying (1) by $\cos\theta$ and (2) by $\sin\theta$ and adding, we get

$$dr = \cos\theta dx + \sin\theta dy \qquad (3)$$

Multiplying (1) by $\sin\theta$ and (2) by $\cos\theta$ and subtracting, we get

$$d\theta = \frac{\cos\theta}{r}dy - \frac{\sin\theta}{r}dx \qquad (4)$$

From (3) and (4), we get

$$\frac{\partial r}{\partial x} = \cos\theta, \qquad \frac{\partial r}{\partial y} = \sin\theta$$

$$\frac{\partial \theta}{\partial y} = \frac{\cos\theta}{r}, \qquad \frac{\partial \theta}{\partial x} = -\frac{\sin\theta}{r}$$

Now, $$\frac{\partial z}{\partial x} = \frac{\partial z}{\partial r}\frac{\partial r}{\partial x} + \frac{\partial z}{\partial \theta}\frac{\partial \theta}{\partial x}$$

$$= \cos\theta\frac{\partial z}{\partial r} - \frac{\sin\theta}{r}\frac{\partial z}{\partial \theta} \qquad (5)$$

$$\left(\frac{\partial z}{\partial x}\right)^2 = \left(\cos\theta\frac{\partial z}{\partial r} - \frac{\sin\theta}{r}\frac{\partial z}{\partial \theta}\right)^2$$

$$= \cos^2\theta\left(\frac{\partial z}{\partial r}\right)^2 + \frac{\sin^2\theta}{r^2}\left(\frac{\partial z}{\partial \theta}\right)^2 - \frac{2\sin\theta\cos\theta}{r}\frac{\partial^2 z}{\partial r\partial \theta} \qquad (6)$$

Similarly $$\frac{\partial z}{\partial y} = \frac{\partial z}{\partial r}\frac{\partial r}{\partial y} + \frac{\partial z}{\partial \theta}\frac{\partial \theta}{\partial y}$$

$$= \sin\theta\frac{\partial z}{\partial r} + \frac{\cos\theta}{r}\frac{\partial z}{\partial \theta} \qquad (7)$$

$$\left(\frac{\partial z}{\partial y}\right)^2 = \sin^2\theta\left(\frac{\partial z}{\partial r}\right)^2 + \frac{\cos^2\theta}{r^2}\left(\frac{\partial z}{\partial \theta}\right)^2 + \frac{2\sin\theta\cos\theta}{r}\frac{\partial^2 z}{\partial r\partial \theta} \qquad (8)$$

Adding (6) and (8), we get

$$\left(\frac{\partial z}{\partial x}\right)^2+\left(\frac{\partial z}{\partial y}\right)^2=\left(\frac{\partial z}{\partial r}\right)^2+\frac{1}{r^2}\left(\frac{\partial z}{\partial \theta}\right)^2$$

Multiplying (a^2-r^2) on both sides,

$$(a^2-r^2)\left\{\left(\frac{\partial z}{\partial x}\right)^2+\left(\frac{\partial z}{\partial y}\right)^2\right\}=(a^2-r^2)\left\{\left(\frac{\partial z}{\partial r}\right)^2+\frac{1}{r^2}\left(\frac{\partial z}{\partial \theta}\right)^2\right\} \qquad (9)$$

Multiplying (5) by x and (7) by y and adding,

$$x\frac{\partial z}{\partial x}+y\frac{\partial z}{\partial y}=\left(x\cos\theta+y\sin\theta\right)\frac{\partial z}{\partial r}+\frac{1}{r}\left(y\cos\theta-x\sin\theta\right)\frac{\partial z}{\partial \theta}$$

$$=\left(r\cos^2\theta+r\sin^2\theta\right)\frac{\partial z}{\partial r}+\frac{r}{r}\left(\sin\theta\cos\theta-\cos\theta\sin\theta\right)\frac{\partial z}{\partial \theta}$$

$$=r\frac{\partial z}{\partial r}$$

Squaring on both sides,

$$\left(x\frac{\partial z}{\partial x}+y\frac{\partial z}{\partial y}\right)^2=r^2\left(\frac{\partial z}{\partial r}\right)^2 \qquad (10)$$

Adding (9) and (10), we get required result

$$\left(x\frac{\partial z}{\partial x}+y\frac{\partial z}{\partial y}\right)^2+(a^2-r^2)\left\{\left(\frac{\partial z}{\partial x}\right)^2+\left(\frac{\partial z}{\partial y}\right)^2\right\}=a^2\left\{\left(\frac{\partial z}{\partial r}\right)^2+\frac{1}{r^2}\left(\frac{\partial z}{\partial \theta}\right)^2\right\}-\left(\frac{\partial z}{\partial \theta}\right)^2$$

$$\left(\because r^2=x^2+y^2\right)$$

Example 3. If $x=r\cos\theta$, $y=r\sin\theta$ then prove that

$$\left(x^2-y^2\right)\left(\frac{\partial^2 u}{\partial x^2}-\frac{\partial^2 u}{\partial y^2}\right)+4xy\frac{\partial^2 u}{\partial x\partial y}=r^2\frac{\partial^2 u}{\partial r^2}-r\frac{\partial u}{\partial r}-\frac{\partial^2 u}{\partial \theta^2}$$

where u is any twice differentiable function of x and y.

Solution. Here, $x=x(r,\theta)$

$$dx=\frac{\partial x}{\partial r}dr+\frac{\partial x}{\partial \theta}d\theta$$

$$=\cos\theta dr-r\sin\theta d\theta \qquad (1)$$

Since $y=y(r,\theta)$

$$dy = \frac{\partial y}{\partial r}dr + \frac{\partial y}{\partial \theta}d\theta$$

$$= \sin\theta dr + r\cos\theta d\theta \tag{2}$$

Multiplying (1) by $\cos\theta$ and (2) by $\sin\theta$ and adding, we get

$$dr = \cos\theta dx + \sin\theta dy \tag{3}$$

Multiplying (1) by $\sin\theta$ and (2) by $\cos\theta$ and subtracting, we get

$$d\theta = -\frac{\sin\theta}{r}dx + \frac{\cos\theta}{r}dy \tag{4}$$

From (3) and (4), we get

$$\frac{\partial r}{\partial x} = \cos\theta, \qquad \frac{\partial r}{\partial y} = \sin\theta$$

$$\frac{\partial \theta}{\partial x} = -\frac{\sin\theta}{r}, \qquad \frac{\partial \theta}{\partial y} = \frac{\cos\theta}{r}$$

Now, $$\frac{\partial u}{\partial x} = \frac{\partial u}{\partial r}\frac{\partial r}{\partial x} + \frac{\partial u}{\partial \theta}\frac{\partial \theta}{\partial x}$$

$$= \frac{\partial u}{\partial r}\cos\theta - \frac{\partial u}{\partial \theta}\frac{\sin\theta}{r}$$

$$= \left(\cos\theta\frac{\partial}{\partial r} - \frac{\sin\theta}{r}\frac{\partial}{\partial \theta}\right)u \tag{5}$$

Similarly $$\frac{\partial u}{\partial y} = \left(\sin\theta\frac{\partial}{\partial r} + \frac{\cos\theta}{r}\frac{\partial}{\partial \theta}\right)u \tag{6}$$

Now, $$\frac{\partial^2 u}{\partial x^2} = \frac{\partial}{\partial x}\left(\frac{\partial u}{\partial x}\right)$$

$$= \left(\cos\theta\frac{\partial}{\partial r} - \frac{\sin\theta}{r}\frac{\partial}{\partial \theta}\right)\left(\cos\theta\frac{\partial u}{\partial r} - \frac{\sin\theta}{r}\frac{\partial u}{\partial \theta}\right)$$

$$= \cos^2\theta\frac{\partial^2 u}{\partial r^2} - \frac{\sin\theta\cos\theta}{r}\frac{\partial^2 u}{\partial r\partial \theta} + \frac{\sin\theta\cos\theta}{r^2}\frac{\partial u}{\partial \theta} - \frac{\sin\theta\cos\theta}{r}\frac{\partial^2 u}{\partial r\partial \theta}$$

$$+ \frac{\sin^2\theta}{r}\frac{\partial u}{\partial r} + \frac{\sin\theta\cos\theta}{r^2}\frac{\partial u}{\partial \theta} + \frac{\sin^2\theta}{r^2}\frac{\partial^2 u}{\partial \theta^2} \tag{7}$$

Similarly $$\frac{\partial^2 u}{\partial y^2} = \sin^2\theta\frac{\partial^2 u}{\partial r^2} + \frac{\sin\theta\cos\theta}{r}\frac{\partial^2 u}{\partial r\partial \theta} - \frac{\sin\theta\cos\theta}{r^2}\frac{\partial u}{\partial \theta} + \frac{\cos^2\theta}{r}\frac{\partial u}{\partial r}$$

$$+\frac{\sin\theta\cos\theta}{r}\frac{\partial^2 u}{\partial r\partial\theta}-\frac{\cos\theta\sin\theta}{r^2}\frac{\partial u}{\partial\theta}+\frac{\cos^2\theta}{r^2}\frac{\partial^2 u}{\partial\theta^2} \qquad (8)$$

Subtracting (8) from (7), we get

$$\left(\frac{\partial^2 u}{\partial x^2}-\frac{\partial^2 u}{\partial y^2}\right)=\cos 2\theta\frac{\partial^2 u}{\partial r^2}-\frac{\sin 2\theta}{r}\frac{\partial^2 u}{\partial r\partial\theta}+\frac{\sin 2\theta}{r^2}\frac{\partial u}{\partial\theta}-\frac{\sin 2\theta}{r}\frac{\partial^2 u}{\partial r\partial\theta}$$

$$-\frac{\cos 2\theta}{r}\frac{\partial u}{\partial r}+\frac{\sin 2\theta}{r^2}\frac{\partial u}{\partial\theta}-\frac{\cos 2\theta}{r^2}\frac{\partial^2 u}{\partial\theta^2}$$

and we have

$$\left(x^2-y^2\right)=r^2\left(\cos^2\theta-\sin^2\theta\right)=r^2\cos 2\theta$$

$$\left(x^2-y^2\right)\left(\frac{\partial^2 u}{\partial x^2}-\frac{\partial^2 u}{\partial y^2}\right)=r^2\cos 2\theta\left\{\cos 2\theta\left(\frac{\partial^2 u}{\partial r^2}-\frac{\partial^2 u}{\partial\theta^2}\right)-\frac{2\sin 2\theta}{r}\frac{\partial^2 u}{\partial r\partial\theta}+\frac{2\sin 2\theta}{r^2}\frac{\partial u}{\partial\theta}-\frac{\cos 2\theta}{r}\frac{\partial u}{\partial r}\right\}$$

$$(9)$$

Now, $$\frac{\partial^2 u}{\partial x\partial y}=\frac{\partial}{\partial x}\left(\frac{\partial u}{\partial y}\right)$$

$$=\left(\cos\theta\frac{\partial}{\partial r}-\frac{\sin\theta}{r}\frac{\partial}{\partial\theta}\right)\left(\sin\theta\frac{\partial u}{\partial r}+\frac{\cos\theta}{r}\frac{\partial u}{\partial\theta}\right)$$

$$=\cos\theta\sin\theta\frac{\partial^2 u}{\partial r^2}+\frac{\cos^2\theta}{r}\frac{\partial^2 u}{\partial r\partial\theta}-\frac{\cos^2\theta}{r^2}\frac{\partial u}{\partial\theta}-\frac{\sin\theta\cos\theta}{r}\frac{\partial u}{\partial r}$$

$$-\frac{\sin^2\theta}{r}\frac{\partial^2 u}{\partial\theta\partial r}-\frac{\sin\theta\cos\theta}{r^2}\frac{\partial^2 u}{\partial\theta^2}+\frac{\sin^2\theta}{r^2}\frac{\partial u}{\partial\theta}$$

$$4xy\frac{\partial^2 u}{\partial x\partial y}=2r^2\sin 2\theta\left(\sin\theta\cos\theta\frac{\partial^2 u}{\partial r^2}+\frac{1}{r}\cos 2\theta\frac{\partial^2 u}{\partial r\partial\theta}-\frac{\cos^2\theta}{r^2}\frac{\partial u}{\partial\theta}\right.$$

$$\left.-\frac{\sin\theta\cos\theta}{r}\frac{\partial u}{\partial r}-\frac{\sin\theta\cos\theta}{r^2}\frac{\partial^2 u}{\partial\theta^2}+\frac{\sin^2\theta}{r^2}\frac{\partial u}{\partial\theta}\right) \qquad (10)$$

Adding (9) and (10), we get required result.

Example 4. If $x=r\cos\theta$, $y=r\sin\theta$, then show that

$$\frac{\partial^2\theta}{\partial x\partial y}=r^{-2}\cos 2\theta.$$

Solution. Here, $$x=x(r,\theta)$$

$$dx=\frac{\partial x}{\partial r}dr+\frac{\partial x}{\partial\theta}d\theta$$

$$= \cos\theta dr - r\sin\theta d\theta \qquad (1)$$

Since $$y = y(r,\theta)$$

$$dy = \frac{\partial y}{\partial r}dr + \frac{\partial y}{\partial \theta}d\theta$$

$$= \sin\theta dr + r\cos\theta d\theta \qquad (2)$$

Multiplying (1) by $\cos\theta$ and (2) by $\sin\theta$ and adding, we get

$$dr = \cos\theta dx + \sin\theta dy \qquad (3)$$

Multiplying (1) by $\sin\theta$ and (2) by $\cos\theta$ and subtracting, we get

$$d\theta = -\frac{\sin\theta}{r}dx + \frac{\cos\theta}{r}dy \qquad (4)$$

From (3) and (4), we get

$$\frac{\partial r}{\partial x} = \cos\theta, \qquad \frac{\partial r}{\partial y} = \sin\theta$$

$$\frac{\partial \theta}{\partial x} = -\frac{\sin\theta}{r}, \qquad \frac{\partial \theta}{\partial y} = \frac{\cos\theta}{r}$$

From (4), $rd\theta = -\sin\theta dx + \cos\theta dy.$

Now differentiating, we get

$$drd\theta + rd^2\theta = -\cos\theta d\theta dx - \sin\theta d\theta dy$$

$$= -(\cos\theta dx + \sin\theta dy)d\theta$$

$$rd^2\theta = -(\cos\theta dx + \sin\theta dy)d\theta - drd\theta$$

$$= -(\cos\theta dx + \sin\theta dy)d\theta - (\cos\theta dx + \sin\theta dy)d\theta$$

$$= -2(\cos\theta dx + \sin\theta dy)d\theta$$

$$= -2(\cos\theta dx + \sin\theta dy)\left(-\frac{\sin\theta}{r}dx + \frac{\cos\theta}{r}dy\right)$$

$$d^2\theta = -\frac{2}{r^2}(\cos\theta dx + \sin\theta dy)(-\sin\theta dx + \cos\theta dy)$$

$$= -\frac{2}{r^2}\left(-\sin\theta\cos\theta dx^2 + \cos 2\theta dxdy + \sin\theta\cos\theta dy^2\right).$$

As $d^2\theta = \frac{\partial^2\theta}{\partial x^2}dx^2 + 2\frac{\partial^2\theta}{\partial x\partial y}dxdy + \frac{\partial^2\theta}{\partial y^2}dy^2$, so we get

$$\frac{\partial^2\theta}{\partial x\partial y} = -r^{-2}\cos 2\theta.$$

4.6 Extreme Values of Explicit Functions

We now investigate the theory of extreme values for explicit functions of more than one variable.

Definition 1. Let $u = f(x, y)$ be the equation which defines u as a function of two independent variables x and y. Then, the function $u = f(x, y)$ has an extreme value at the point (a,b) if the increment $\Delta f = f(a+h, b+k) - f(a,b)$ preserves the same sign for all values of h and k such that $|h| < \delta$, $|k| < \delta$ where δ is a sufficiently small positive number. If Δf is negative then the value is maximum and if Δf is positive then the value is minimum.

__Necessary condition for extreme value__

The necessary condition that $f(a,b)$ should be an extreme value is that both $f_x(a,b)$ and $f_y(a,b)$ are zero. Values of (x, y) at which $df = 0$ are called stationary values.

Or A necessary condition for $f(x, y)$ to have an extreme value at (a,b) is that $f_x(a,b) = 0$, $f_y(a,b) = 0$ provided that these partial derivatives exist.

If $f(a,b)$ is an extreme value of the function $f(x, y)$ of two variables then it must also be an extreme value of both the functions $f(x,b)$ and $f(a, y)$ of one variable.

But the necessary condition that these have extreme values at $x = a$ and $y = b$ respectively is $f_x(a,b) = 0$ and $f_y(a,b) = 0$.

__Sufficient condition for extreme value__

The value $f(a,b)$ is an extreme value of $f(x, y)$ if $f_x(a,b) = 0$, $f_y(a,b) = 0$ and also $f_{xx}.f_{yy} > \left(f_{xy}\right)^2$ and the value is maximum or minimum according as f_{xx} or f_{yy} is negative or positive respectively.

Here, $A = f_{xx}$, $C = f_{yy}$, $B = f_{xy}$

(i) If $AC - B^2 > 0$, then $f(a,b)$ is a maximum value if $A < 0$ and a minimum value if $A > 0$.

(ii) If $AC - B^2 < 0$, then $f(a,b)$ is not an extreme value.

(iii) If $AC - B^2 = 0$, this is doubtful case, in which the sign of $f(a+h, b+k) - f(a,b)$ depends on h and k and requires further investigation.

Example 1. Find the extreme value of the function $f(x, y) = x^2 - xy + y^2 + 3x - 2y + 1$.

Solution. Here, $f(x, y) = x^2 - xy + y^2 + 3x - 2y + 1$

$\therefore f_x = 2x - y + 3$, $\quad f_{xx} = 2$

$$f_y = -x + 2y - 2, \qquad f_{yy} = 2, \qquad f_{xy} = -1.$$

For extreme values, $f_x = 0$, $f_y = 0$

$$\therefore 2x - y = -3 \text{ and } -x + 2y = 2$$

$$\therefore y = \frac{1}{3},\ x = -\frac{4}{3}$$

Thus, the extreme point is $\left(-\frac{4}{3}, \frac{1}{3}\right)$.

At $\left(-\frac{4}{3}, \frac{1}{3}\right)$, $\qquad A = f_{xx} = 2,\ B = f_{xy} = -1,\ C = f_{xy} = 2$

Now, $\qquad AC - B^2 = 4 - 1 = 3 > 0$ and $A = 2 > 0$

$\therefore \left(-\frac{4}{3}, \frac{1}{3}\right)$ is a point of minimum and minimum value $= f\left(-\frac{4}{3}, \frac{1}{3}\right)$

$$= \frac{16}{9} + \frac{4}{9} + \frac{1}{9} - 4 - \frac{2}{3} + 1 = -\frac{4}{3}.$$

Example 2. Show that $f(x, y) = 2x^4 - 3x^2y + y^2$ has neither maximum nor minimum at (0, 0).

Solution. Here, $f(x, y) = 2x^4 - 3x^2y + y^2$

$$\therefore f_x = 8x^3 - 6xy, \qquad f_{xx} = 24x^2 - 6y$$

$$f_y = -3x^2 + 2y, \qquad f_{yy} = 2, \qquad f_{xy} = -6x$$

For extreme values, $f_x = 0$, $f_y = 0$

$$8x^3 - 6xy = 0 \text{ and } -3x^2 + 2y = 0$$

$$2x\left(4x^2 - 3y\right) = 0 \text{ and } y = \frac{3x^2}{2}$$

$$\therefore x = 0 \text{ or } y = \frac{4x^2}{3}$$

If $x = 0 \Rightarrow y = 0$

If $y = \frac{4x^2}{3}$ and $y = \frac{3x^2}{2} \Rightarrow \frac{4x^2}{3} = \frac{3x^2}{2}$, which is not possible.

So, stationary point is (0,0).

Now, $\quad A = f_{xx}(0,0) = 0,\ B = f_{xy}(0,0) = 0,\ C = f_{yy}(0,0) = 2$

$$\therefore AC - B^2 = 0 - 0 = 0$$

So, doubtful case and further investigation is required.

Now, $\Delta f = f(0+h, 0+k) - f(0,0)$

$= f(h,k) - f(0,0)$

$= 2h^4 - 3h^2k + k^2 = (2h^2 - k)(h^2 - k)$

If $h^2 - k > 0$ and $2h^2 - k > 0$ i.e. $h^2 > k$ and $h^2 > \frac{k}{2}$ then $\Delta f > 0$.

If $h^2 - k < 0$ and $2h^2 - k > 0$ i.e. $h^2 < k$ and $h^2 > \frac{k}{2}$ then $\Delta f < 0$.

So, for different values of h and k, Δf does not have the same sign. Hence, f has neither maximum nor minimum at (0, 0).

Example 3. Find the extreme value of $x^3 - 3axy + y^3; a > 0.$.

Solution. Here, $f(x, y) = x^3 - 3axy + y^3$

$\therefore f_x = 3x^2 - 3ay$, $\quad f_{xx} = 6x$

$f_y = 3y^2 - 3ax$, $\quad f_{yy} = 6y$, $\quad f_{xy} = -3a$

For extreme value, we put $f_x = 0$, $f_y = 0$

$$3x^2 - 3ay = 0 \text{ and } 3y^2 - 3ax = 0$$

$$y = \frac{x^2}{a} \text{ and } x = \frac{y^2}{a}$$

After solving, the stationary points are (0, 0) and (a, a).

Now, $A = f_{xx}(0,0) = 0$, $B = f_{xy}(0,0) = -3a$, $C = f_{yy}(0,0) = 0$

$$AC - B^2 = -9a^2 < 0 \text{ at } (0,0).$$

So, given function has no extreme value at (0,0).

Now, $A = f_{xx}(a,a) = 6a$, $B = f_{xy}(a,a) = -3a$, $C = f_{yy}(a,a) = 6a$

$$AC - B^2 = 36a^2 - 9a^2 = 27a^2 > 0 \text{ at } (a,a).$$

$$\& \; A = 6a > 0$$

Hence, the given function has minimum value at (a, a).

Example 4. Let $u = xy + \frac{a^3}{x} + \frac{a^3}{y}$,

$$\frac{\partial u}{\partial x} = y - \frac{a^3}{x^2}$$
$$\frac{\partial u}{\partial y} = x - \frac{a^3}{y^2}.$$

Putting x= a, y = a

Hence $\frac{\partial^2 u}{\partial x^2} = \frac{2a^3}{x^3} = 2,\ \frac{\partial^2 u}{\partial x \partial y} = 1,\ \frac{\partial^2 u}{\partial y^2} = \frac{2a^3}{y^3} = 2.$

Therefore r and t are positive when x = a = y and rt- s^2 = 2.2-1= 3 (positive). Therefore, there is a minimum value of u viz. u = $3a^2$.

Example 5. Let

$$f(x, y) = y^2 + x^2y + x^4.$$

It can be verified that

$$f_x(0,0) = 0,\ f_y(0,0) = 0$$
$$f_{xx}(0,0) = 0,\ f_{yy}(0,0) = 2$$
$$f_{xy}(0,0) = 0.$$

So at the origin, we have

$$f_{xx}f_{yy} = {f_{xy}}^2.$$

However, on writing

$$y^2 + x^2y + x^4 = (y + \tfrac{1}{2}x^2)^2 + \tfrac{3x^4}{4}.$$

It is clear that f(x, y) has a minimum value at the origin, since

$$\Delta f = f(h, k) - f(0,0) = (k + \tfrac{h^2}{2})^2 + \tfrac{3h^4}{4}$$

is greater than zero for all values of h and k.

4.7 Stationary Values of Implicit Functions

To find the stationary values of the function

$$f(x_1, x_2, \ldots\ldots\ldots, x_n, u_1, u_2, \ldots\ldots\ldots\ldots\ldots, u_m) \quad (1)$$

of $(n+m)$ variables which are connected by m differentiable equations

$$\varphi_r(x_1, x_2, \ldots\ldots\ldots\ldots, x_n, u_1, u_2, \ldots\ldots\ldots\ldots\ldots\ldots, u_m) = 0; r = 1, 2, \ldots\ldots\ldots\ldots\ldots, m \qquad (2)$$

If the m variables $u_1, u_2, \ldots\ldots\ldots\ldots, u_m$ are determinate as functions of $x_1, x_2, \ldots\ldots\ldots\ldots, x_n$ from the system of m equations of (2), then f can be regarded as a function of n independent variables $x_1, x_2, \ldots\ldots\ldots, x_n$.

At a stationary point of f, $df = 0$.

Hence at a stationary point, $0 = df = f_{x_1} dx_1 + f_{x_2} dx_2 + \ldots\ldots\ldots + f_{x_n} dx_n + f_{u_1} du_1 + \ldots\ldots\ldots\ldots + f_{u_m} du_m$
(3)

Again differentiating the equation (2), we get

$$\left.\begin{array}{l}
\frac{\partial \phi_1}{\partial x_1} dx_1 + \ldots\ldots\ldots + \frac{\partial \phi_1}{\partial x_n} dx_n + \frac{\partial \phi_1}{\partial u_1} du_1 + \ldots\ldots\ldots\ldots + \frac{\partial \phi_1}{\partial u_m} du_m = 0 \\
\frac{\partial \phi_2}{\partial x_1} dx_1 + \ldots\ldots\ldots + \frac{\partial \phi_2}{\partial x_n} dx_n + \frac{\partial \phi_2}{\partial u_1} du_1 + \ldots\ldots\ldots + \frac{\partial \phi_2}{\partial u_m} du_m = 0 \\
\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
\frac{\partial \phi_m}{\partial x_1} dx_1 + \ldots\ldots\ldots + \frac{\partial \phi_m}{\partial x_n} dx_n + \frac{\partial \phi_m}{\partial u_1} du_1 + \ldots\ldots\ldots + \frac{\partial \phi_m}{\partial u_m} du_m = 0
\end{array}\right\} \qquad (4)$$

From these m equations of (4), the differentials $du_1, du_2, \ldots\ldots\ldots\ldots\ldots, du_m$ of the m dependent variables may be found in terms of the n differentials $dx_1, dx_2, \ldots\ldots\ldots, dx_n$ and are substituted in (3). This way df has been expressed in terms of the differentials of the independent variables, and since the differentials of the independent variables are arbitrary and $df = 0$, the coefficients of each of these n differentials may be equated to zero. These n equations together with the m equations of (2) constitute a system of $(n+m)$ equations to determine the $(n+m)$ coordinates of the stationary points of f.

Example 1. $F(x, y, z)$ is a function subject to the constraint condition $G(x, y, z) = 0$. Show that at a stationary point.

$$F_x G_y - F_y G_x = 0.$$

Solution. We may consider z as a function of the independent variables x, y.

At a stationary point, $dF = 0$

$$\therefore \qquad 0 = dF = F_x dx + F_y dy + F_z dz. \qquad (1)$$

Differentiating the relation $G(x, y, z) = 0$, we get

$$G_x dx + G_y dy + G_z dz = 0. \qquad (2)$$

Putting the values of dz from (2) into (1), or what is same thing, eliminating dz from (1) and (2), we get

$$(F_x G_z - G_x F_z)dx + (F_y G_z - G_y F_z)dy = 0$$

Since dx, dy (being differentials of independent variables) are arbitrary, therefore

$$F_x G_z - G_x F_z = 0$$
$$F_y G_z - G_y F_z = 0$$

which gives $\quad F_x G_y - G_x F_y = 0.$

4.8 Lagrange Multipliers Method

In this method, we discuss the determination of stationary points from a modified point of view. This process consists in the introduction of undetermined multipliers, a method due to Lagrange. After his name, this method also called Lagrange's method of undetermined multipliers.

Let $u = \phi(\mathrm{x}_1, \mathrm{x}_2, \dots, \mathrm{x}_\mathrm{n})$ be a function of n variables which are connected by m equations

$$f_1(\mathrm{x}_1, \mathrm{x}_2, \dots, \mathrm{x}_\mathrm{n}) = 0, f_2(\mathrm{x}_1, \mathrm{x}_2, \dots, \mathrm{x}_\mathrm{n}) = 0, \dots, f_m(\mathrm{x}_1, \mathrm{x}_2, \dots, \mathrm{x}_\mathrm{n}) = 0,$$

So that only n-m variables are independent.

When u is maximum or minimum

$$du = \frac{\partial u}{\partial \mathrm{x}_1}\, d\mathrm{x}_1 + \frac{\partial u}{\partial \mathrm{x}_2}\, d\mathrm{x}_2 + \frac{\partial u}{\partial \mathrm{x}_3}\, d\mathrm{x}_3 + \cdots + \frac{\partial u}{\partial \mathrm{x}_\mathrm{n}}\, d\mathrm{x}_\mathrm{n} = 0$$

Also $$df_1 = \frac{\partial f_1}{\partial \mathrm{x}_1}\, d\mathrm{x}_1 + \frac{\partial f_1}{\partial \mathrm{x}_2}\, d\mathrm{x}_2 + \frac{\partial f_1}{\partial \mathrm{x}_3}\, d\mathrm{x}_3 + \cdots + \frac{\partial f_1}{\partial \mathrm{x}_\mathrm{n}}\, d\mathrm{x}_\mathrm{n} = 0$$

$$df_2 = \frac{\partial f_2}{\partial \mathrm{x}_1}\, d\mathrm{x}_1 + \frac{\partial f_2}{\partial \mathrm{x}_2}\, d\mathrm{x}_2 + \frac{\partial f_2}{\partial \mathrm{x}_3}\, d\mathrm{x}_3 + \cdots + \frac{\partial f_2}{\partial \mathrm{x}_\mathrm{n}}\, d\mathrm{x}_\mathrm{n} = 0$$

……………………………………………………………………

……………………………………………………………………..

$$df_m = \frac{\partial f_m}{\partial \mathrm{x}_1}\, d\mathrm{x}_1 + \frac{\partial f_m}{\partial \mathrm{x}_2}\, d\mathrm{x}_2 + \frac{\partial f_m}{\partial \mathrm{x}_3}\, d\mathrm{x}_3 + \cdots + \frac{\partial f_m}{\partial \mathrm{x}_\mathrm{n}}\, d\mathrm{x}_\mathrm{n} = 0$$

Multiplying all these respectively by 1, $\lambda_1, \lambda_2, \dots, \lambda_n$ and adding, we get a result which may be written

$$P_1\, d\mathrm{x}_1 + P_2\, d\mathrm{x}_2 + P_3\, d\mathrm{x}_3 + \cdots + P_n d\mathrm{x}_\mathrm{n} = 0,$$

Where $P_r = \frac{\partial u}{\partial \mathrm{x}_\mathrm{r}} + \lambda_1 \frac{\partial f_1}{\partial \mathrm{x}_\mathrm{r}} + \lambda_2 \frac{\partial f_2}{\partial \mathrm{x}_\mathrm{r}} + \cdots + \lambda_m \frac{\partial f_m}{\partial \mathrm{x}_\mathrm{r}}$

The m quantities $\lambda_1, \lambda_2, \dots, \lambda_m$ are of our choice. Let us choose them so as to satisfy the m linear equations

$$\mathrm{P}_1 = \mathrm{P}_2 = \dots\dots\dots = \mathrm{P}_\mathrm{m}.$$

The above equation is now reduced to

$$P_{m+1} dx_{m+1} + P_{m+2} dx_{m+2} + \cdots + P_n dx_n = 0$$

It is indifferent which n-m of the n variables are regarded as independent. Let them be $x_{m+1}, x_{m+2}, \ldots, x_n$. Then since n-m quantities $dx_{m+1}, dx_{m+2}, \ldots dx_n$ are all independent, their coefficients must be separately zero. Thus we obtain the additional n-m equations

$$P_{m+1} = P_{m+2} = \ldots\ldots\ldots\ldots = P_n = 0.$$

Thus the m+n equations $f_1 = f_2 = \ldots\ldots\ldots\ldots = f_m = 0$ and $P_1 = P_2 = \ldots\ldots\ldots\ldots = P_n = 0$ determine the m multipliers $\lambda_1, \lambda_2, \ldots, \lambda_m$ and values of n variables $x_1, x_2, \ldots, x_n$ for which maximum and minimum values of u are possible.

Example 1. Find the length of the axes of the section of the ellipsoid $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$ by the plane $lx + my + nz = 0$.

Solution. We have to find the extreme values of the function r^2 where $r^2 = x^2 + y^2 + z^2$, subject to the equations of the condition

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} - 1 = 0,$$

$$lx + my + nz = 0.$$

Then
$$xdx + ydy + zdz = 0 \tag{1}$$

$$\frac{x}{a^2}dx + \frac{y}{b^2}dy + \frac{z}{c^2}dz = 0, \tag{2}$$

$$ldx + mdy + ndz = 0 \tag{3}$$

Multiplying these equations by 1, λ_1, λ_2 and adding we get

$$x + \lambda_1 \frac{x}{a^2} + \lambda_2 l = 0 \tag{4}$$

$$y + \lambda_1 \frac{y}{b^2} + \lambda_2 m = 0 \tag{5}$$

$$z + \lambda_1 \frac{z}{c^2} + \lambda_2 n = 0 \tag{6}$$

Multiplying (4), (5) and (6) by x, y, z and adding we get

$$(x^2 + y^2 + z^2) + \lambda_1 \left(\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2}\right) + \lambda_2(lx + my + nz) = 0$$

or
$$r^2 + \lambda_1 = 0 \Longrightarrow \lambda_1 = -r^2.$$

From (4), (5) and (6), we have

$$x = \frac{\lambda_2 l}{(\frac{r^2}{a^2} - 1)}, y = \frac{\lambda_2 m}{(\frac{r^2}{b^2} - 1)}, z = \frac{\lambda_2 n}{(\frac{r^2}{c^2} - 1)}$$

But $lx + my + nz = 0 \Rightarrow \lambda_2\left(\frac{l^2a^2}{r^2-a^2}+\frac{m^2b^2}{r^2-b^2}+\frac{n^2c^2}{r^2-c^2}\right)=0$ and since $\lambda_2 \neq 0$ the equation giving the values of r^2, which are the squares of the length of semi-axes required (quadratic in r^2) is $\left(\frac{l^2a^2}{r^2-a^2}+\frac{m^2b^2}{r^2-b^2}+\frac{n^2c^2}{r^2-c^2}\right)=0$.

Example 2. Investigate the maximum and minimum radii vector of the sector of "surface of elasticity" $(x^2+y^2+z^2)^2=a^2x^2+y^2b^2+z^2c^2$ made by the plane $lx + my + nz = 0$.

Solution. We have

$$xdx + ydy + zdz = 0 \qquad (1)$$

$$a^2xdx + b^2ydy + c^2zdz = 0 \qquad (2)$$

$$ldx + mdy + ndz = 0 \qquad (3)$$

Multiplying these equations by 1, $\lambda_1\lambda_2$ and adding we get

$$x + a^2x\lambda_1 + l\lambda_2 = 0 \qquad (4)$$

$$y + b^2y\lambda_1 + m\lambda_2 = 0 \qquad (5)$$

$$z + c^2z\lambda_1 + n\lambda_2 = 0 \qquad (6)$$

Multiplying(4), (5) and (6) by x, y, z respectively and adding we get

$$\left(x^2+y^2+z^2\right)+\lambda_1\left(a^2x^2+y^2b^2+z^2c^2\right)+\lambda_2\left(lx+my+nz\right)=0$$

$$\Rightarrow r^2+\lambda_1 r^4=0 \quad \Rightarrow \lambda_1=-\frac{1}{r^2}$$

$$\Rightarrow x=\frac{\lambda_2 lr^2}{a^2-r^2},\quad y=\frac{\lambda_2 mr^2}{b^2-r^2},\quad z=\frac{\lambda_2 nr^2}{c^2-r^2}.$$

Then $lx + my + nz = 0 \Rightarrow \frac{\lambda_2 l^2r^2}{a^2-r^2}+\frac{\lambda_2 m^2r^2}{b^2-r^2}+\frac{\lambda_2 n^2r^2}{c^2-r^2}=0.$

$$\Rightarrow \frac{l^2}{r^2-a^2}+\frac{m^2}{r^2-b^2}+\frac{n^2}{r^2-c^2}$$

It is quadratic in r^2 and give its required values.

Example 3. Prove that the volume of the greatest rectangular parallelepiped that can be inscribed in the ellipsoid $\frac{\partial(\xi,\eta,\zeta)}{\partial(\alpha,\beta,\gamma)}=\begin{vmatrix}-1 & -1 & -1\\ \beta+\gamma & \gamma+\alpha & \alpha+\beta\\ -\beta\gamma & -\gamma\alpha & -\alpha\beta\end{vmatrix}$.

Solution. Volume of the parallelepiped $= 8xyz$. Its maximum value is to find under the condition that it is inscribed in the ellipsoid $\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}=1$, we have

$$u = 8xyz$$

$$f_1 = \frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}=1$$

Therefore

$$du = 8yzdx + 8xzdy + 8xydz = 0 \qquad (1)$$

$$df_1 = \frac{2x}{a^2}dx+\frac{2y}{b^2}dy+\frac{2z}{c^2}dz=0 \qquad (2)$$

Multiplying (1) by 1 and (2) by λ and adding we get

$$yz+\frac{x}{a^2}\lambda=0 \qquad (3)$$

$$zx+\frac{y}{b^2}\lambda=0 \qquad (4)$$

$$xy+\frac{z}{c^2}\lambda=0 \qquad (5)$$

From (3), (4) and (5), we get

$$\lambda=-\frac{a^2yz}{x}=-\frac{b^2zx}{y}=-\frac{c^2xy}{z}$$

and so $\frac{a^2yz}{x}=\frac{b^2zx}{y}=\frac{c^2xy}{z}$

Dividing throughout by xyz we get

$$\frac{a^2}{x^2}=\frac{b^2}{y^2}=\frac{c^2}{z^2} \quad \left(\because \frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}=1\right).$$

Hence $\frac{3x^2}{x^2}=1$ or $x=\frac{a}{\sqrt{3}}$ .Similarly $y=\frac{b}{\sqrt{3}}$, $z=\frac{c}{\sqrt{3}}$

It follows therefore that $u=8xyz=\frac{8abc}{3\sqrt{3}}$

Example 4. Find the point of the circle $x^2 + y^2 + z^2 = 1$, $lx+my+nz=0$ at which the function $u = ax^2 + by^2 + cz^2 + 2fyz + 2gzx + 2hxy$ attains its greatest and least value.

Solution. We have

$$u = ax^2 + by^2 + cz^2 + 2fyz + 2gzx + 2hxy$$

$$f_1 = lx+my+nz$$

$$f_2 = x^2 + y^2 + z^2 - 1$$

Then

$axdx + bydy + czdz + fydz + fzdy + ... + gzdx + gxdz + hxdy + hydx$

$$ldx + mdy + ndz = 0$$

$$xdx + ydy + zdz = 0$$

Multiplying these equations by 1, λ_1, λ_2 and adding, we get

$$ax + hy + gz + \lambda_1 l + \lambda_2 x = 0$$

$$by + hx + fz + \lambda_1 m + \lambda_2 y = 0$$

$$cz + gx + fy + \lambda_1 n + \lambda_2 z = 0.$$

Multiplying by x, y, z and adding we get

$$u + \lambda_2 = 0 \quad \Rightarrow \quad \lambda_2 = -u.$$

Putting all the values in the above equation we have

$$x(a-u) + hy + gz + l\lambda_1 = 0$$

$$hx + y(b-u) + fz + m\lambda_1 = 0$$

$$gx + fy + z(c-u) + n\lambda_1 = 0$$

$$lx + my + nz + 0 = 0.$$

Eliminating x, y, z and λ_1 ,we get

$$\begin{vmatrix} a-u & h & g & 1 \\ h & b-u & f & m \\ g & f & c-u & n \\ l & m & n & 0 \end{vmatrix} = 0.$$

Example 5. If a, b, c are positive and

$$u = (a^2x^2 + b^2y^2 + c^2z^2)/_{x^2y^2z^2}, \quad ax^2 + by^2 + cz^2 = 1,$$

Show that a stationary value of u is given by

$$x^2 = \frac{\mu}{2a(\mu+a)}, y^2 = \frac{\mu}{2b(\mu+b)}, z^2 = \frac{\mu}{2c(\mu+c)}$$

where μ is the +ve root of the cubic

$$\mu^3 - (bc + ca + ab)\mu - 2abc = 0.$$

Solution. We have

$$u = \frac{(a^2x^2 + b^2y^2 + c^2z^2)}{x^2y^2z^2} \qquad (1)$$

$$ax^2 + by^2 + cz^2 = 1 \qquad (2)$$

Differentiating (1) we get

$$\sum \frac{1}{x^3}\left(\frac{b^2}{z^2}+\frac{c^2}{y^2}\right)dx = 0$$

which on multiplication with $x^2y^2z^2$ yields

$$\sum \frac{1}{x}\left(b^2y^2+c^2z^2\right)dx = 0 \qquad (3)$$

Differentiating (2) we have

$$\sum axdx = 0 \qquad (4)$$

Using Lagrange's multiplier, we obtain

$$\frac{1}{x}(b^2y^2+c^2z^2) = \mu ax$$

i.e.

$$b^2y^2+c^2z^2 = \mu ax^2 \qquad (5)$$

$$c^2z^2+a^2x^2 = \mu by^2 \qquad (6)$$

$$a^2x^2+b^2y^2 = \mu cz^2 \qquad (7)$$

Then (6) + (7) - (5) yields

$$2a^2x^2 = \mu(by^2+cz^2-ax^2)$$

$$= \mu(1-2ax^2) \qquad \text{(By (2))}$$

Therefore

$$2a(a+\mu)x^2 = \mu$$

$$\Rightarrow x^2 = \frac{\mu}{2a(a+\mu)}.$$

Similarly $y^2 = \frac{\mu}{2b(b+\mu)}$ and $z^2 = \frac{\mu}{2c(c+\mu)}$.

Substituting these values of x^2, y^2, z^2 in (2), we obtain

$$\frac{\mu}{2(a+\mu)}+\frac{\mu}{2(b+\mu)}+\frac{\mu}{2(c+\mu)} = 1$$

which equals to

$$\mu^3-(bc+ca+ab)\mu-2abc = 0. \qquad (8)$$

Since a, b, c are positive, any one of (5), (6), (7) shows that μ must be positive. Hence μ is a positive root of (8).

4.9 Jacobian and its Properties

In this section, we give definition of Jacobian and discuss its properties.

4.9.1 Jacobian

If $u_1, u_2, \ldots, u_n$ be n differentiable functions of the n variables $x_1, x_2, \ldots, x_n$ then the determinant

$$\begin{pmatrix} \frac{\partial u_1}{\partial x_1}, & \frac{\partial u_1}{\partial x_2}, \ldots, & \frac{\partial u_1}{\partial x_n} \\ \frac{\partial u_2}{\partial x_1}, & \frac{\partial u_2}{\partial x_2}, \ldots, & \frac{\partial u_2}{\partial x_n} \\ \frac{\partial u_n}{\partial x_1}, & \frac{\partial u_n}{\partial x_2}, \ldots, & \frac{\partial u_n}{\partial x_n} \end{pmatrix}$$

is called the **Jacobian** of $u_1, u_2, \ldots, u_n$ with regard to $x_1, x_2, \ldots, x_n$. The determinant is often denoted by

$$\frac{\partial(u_1, u_2, \ldots\ldots, u_n)}{\partial(x_1, x_2, \ldots\ldots, x_n)} \quad or \quad J\frac{(u_1, u_2, \ldots\ldots, u_n)}{(x_1, x_2, \ldots\ldots, x_n)}$$

or shortly J, when there can be no doubt as to the variables referred to.

Theorem 1. If $u_1, u_2, \ldots, u_n$ be n differentiable functions of the n independent variables $x_1, x_2, \ldots, x_n$ and there exists an identical differentiable functional relation $\phi(u_1, u_2, \ldots, u_n) = 0$ which does not involve the x's explicitly, then the Jacobian

$$\frac{\partial(u_1, u_2, \ldots, u_n)}{\partial(x_1, x_2, \ldots, x_n)}$$

vanishes identically provided that ϕ as a function of the u's has no stationary values in the domain considered.

Proof. Since

$$\phi(u_1, u_2, \ldots, u_n) = 0.$$

We have

$$\frac{\partial\phi}{\partial u_1}du_1 + \frac{\partial\phi}{\partial u_2}du_2 + \cdots + \frac{\partial\phi}{\partial u_n}du_n = 0 \tag{1}$$

But

$$\begin{cases} du_1 = \frac{\partial u_1}{\partial x_1}dx_1 + \frac{\partial u_1}{\partial x_2}dx_2 + \cdots + \frac{\partial u_1}{\partial x_n}dx_n \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ du_n = \frac{\partial u_n}{\partial x_1}dx_1 + \frac{\partial u_n}{\partial x_2}dx_2 + \cdots + \frac{\partial u_n}{\partial x_n}dx_n \end{cases} \tag{2}$$

On substituting these values in (1) we get an equation of the form

$$A_1 dx_1 + A_2 dx_2 + \cdots + A_n dx_n = 0 \tag{3}$$

And since $dx_1, dx_2, \ldots, dx_n$ are arbitrary differentials of independent variables, it follows that

$A_1 = 0, A_2 = 0, \ldots, A_n = 0$

In other words

$$\begin{cases} \frac{\partial\phi}{\partial u_1}.\frac{\partial u_1}{\partial x_1} + \frac{\partial\phi}{\partial u_2}.\frac{\partial u_2}{\partial x_1} + \cdots \frac{\partial\phi}{\partial u_n}.\frac{\partial u_n}{\partial x_1} = 0 \\ \frac{\partial\phi}{\partial u_1}.\frac{\partial u_1}{\partial x_2} + \frac{\partial\phi}{\partial u_2}.\frac{\partial u_2}{\partial x_2} + \cdots \frac{\partial\phi}{\partial u_n}.\frac{\partial u_n}{\partial x_2} = 0 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \frac{\partial\phi}{\partial u_1}.\frac{\partial u_1}{\partial x_n} + \frac{\partial\phi}{\partial u_2}.\frac{\partial u_2}{\partial x_n} + \cdots \frac{\partial\phi}{\partial u_n}.\frac{\partial u_n}{\partial x_n} = 0 \end{cases} \tag{4}$$

And since by the hypothesis, we cannot have

$$\frac{\partial\phi}{\partial u_1} = \frac{\partial\phi}{\partial u_2} = \cdots = \frac{\partial\phi}{\partial u_n} = 0$$

On eliminating the partial derivatives of ϕ from the set of equation (4) we get

$$\frac{\partial(u_1, u_2, \ldots, u_n)}{\partial(x_1, x_2, \ldots, x_n)} = 0$$

which establishes the theorem.

Theorem 2. If $u_1, u_2, \ldots, u_n$ be n functions of n variables $x_1, x_2, \ldots, x_n$ say $u_m = f_m(x_1, x_2, \ldots, x_m), (m = 1, 2, \ldots, n)$ and if $\frac{\partial(u_1, u_2, \ldots, u_n)}{\partial(x_1, x_2, \ldots, x_n)} = 0$, then if all differential coefficients concerned are continuous, there exists a functional relation connecting some or all of the variables $u_1, u_2, \ldots, u_n$ which is independent of $x_1, x_2, \ldots, x_n$.

Proof. First we prove the theorem when n=2. We have $u = f(x, y), v = g(x, y)$ and

$$\begin{vmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{vmatrix} = 0$$

If v does not depend on y, then $\frac{\partial v}{\partial y} = 0$ and so either $\frac{\partial u}{\partial y} = 0$ or else $\frac{\partial v}{\partial x} = 0$. In the former case u and v are the functions of x only, and the functional relation sought is obtained from

$$u = f(x), v = g(x)$$

By regarding x as a function of v and substituting in $u = f(x)$. In the latter case v is constant, and the functional relation is v = a. If v does depend on y, since $\frac{\partial v}{\partial y} \neq 0$ the equation v = g(x, y) defines y as a function of x and v, say

$$y = \psi(x, v),$$

And on substituting in the other equation we get an equation of the form

$$u = F(x, v).$$

(The function F[x, g(x, y)] is the same function of x and y as f(x, y))

Then

$$0=\begin{vmatrix}\frac{\partial u}{\partial x} & \frac{\partial u}{\partial y}\\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y}\end{vmatrix}=\begin{vmatrix}\frac{\partial F}{\partial x}+\frac{\partial F}{\partial v}\cdot\frac{\partial v}{\partial x} & \frac{\partial F}{\partial v}\cdot\frac{\partial v}{\partial y}\\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y}\end{vmatrix}=\begin{vmatrix}\frac{\partial F}{\partial x} & 0\\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y}\end{vmatrix}$$

(obtained on multiplying the second row by $\frac{\partial F}{\partial u}$ and subtracting from the first) and so, either $\frac{\partial v}{\partial y}=0,$ which is contrary to hypothesis or else $\frac{\partial F}{\partial x}=0,$ so that F is a function of v only; hence the functional relation is

$$u = F(v)$$

Now assume that the theorem holds for $n-1$.

Now u_n must involve one of the variables at least, for if not there is a functional relation $u_n = a$. Let one such variable be called x_n since $\frac{\partial u_n}{\partial x_n} \neq 0$ we can solve the equation

$$u_n = f_n(x_1, x_2, ..., x_n)$$

for x_n in terms of $x_1, x_2, ..., x_{n-1}$ and u_n, and on substituting this value in each of the other equations we get $n-1$ equations of the form

$$u_r = g_r(x_1, x_2, ..., x_{n-1}, u_n),\ (r = 1, 2, ..., n-1) \qquad (1)$$

If now we substitute $f_n(x_1, x_2, ..., x_n)$ for u_n the functions $g_r(x_1, x_2, ..., x_{n-1}, u_n)$ become

$$f_r(x_1, x_2, ..., x_{n-1}, x_n),\ (r = 1, 2, ..., n-1)$$

Then

$$0=\begin{vmatrix}\frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2}, \dots, & \frac{\partial f_1}{\partial x_n}\\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2}, \dots, & \frac{\partial f_2}{\partial x_n}\\ \dots & \dots & \dots\\ \frac{\partial f_n}{\partial x_1} & \frac{\partial f_n}{\partial x_2}, \dots, & \frac{\partial f_n}{\partial x_n}\end{vmatrix}$$

$$= \begin{vmatrix} \frac{\partial g_1}{\partial x_1}+\frac{\partial g_1}{\partial u_n}.\frac{\partial u_n}{\partial x_1},\ldots,\frac{\partial g_1}{\partial x_{n-1}}+\frac{\partial g_1}{\partial u_n}.\frac{\partial u_n}{\partial x_{n-1}},\frac{\partial g_1}{\partial u_n}.\frac{\partial u_n}{\partial x_n} \\ \frac{\partial g_2}{\partial x_1}+\frac{\partial g_2}{\partial u_n}.\frac{\partial u_n}{\partial x_1},\ldots,\frac{\partial g_2}{\partial x_{n-1}}+\frac{\partial g_2}{\partial u_n}.\frac{\partial u_n}{\partial x_{n-1}},\frac{\partial g_2}{\partial u_n}.\frac{\partial u_n}{\partial x_n} \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \frac{\partial u_n}{\partial x_1}, \qquad \ldots, \quad \frac{\partial u_n}{\partial x_{n-1}}, \qquad\qquad \frac{\partial u_n}{\partial x_n} \end{vmatrix}$$

$$= \begin{vmatrix} \frac{\partial g_1}{\partial x_1},\ldots,\frac{\partial g_1}{\partial x_{n-1}}, & 0 \\ \frac{\partial g_2}{\partial x_1},\ldots,\frac{\partial g_2}{\partial x_{n-1}}, & 0 \\ \cdots\cdots\cdots\cdots\cdots & \\ \frac{\partial u_n}{\partial x_1},\ldots,\frac{\partial u_n}{\partial x_{n-1}}, & \frac{\partial u_n}{\partial x_n} \end{vmatrix}$$

by subtracting the elements of the last row multiplied by

$$\frac{\partial g_1}{\partial u_n},\frac{\partial g_2}{\partial u_n}\ldots,\frac{\partial g_n}{\partial u_n}$$

from each of the others. Hence

$$\frac{\partial u_n}{\partial x_n}.\frac{\partial(g_1,g_2,\ldots,g_{n-1})}{\partial(x_1,x_2,\ldots,x_{n-1})}=0.$$

Since $\frac{\partial u_n}{\partial x_n}\neq 0$ we must have $\frac{\partial(g_1,g_2,\ldots,g_{n-1})}{\partial(x_1,x_2,\ldots,x_{n-1})}=0,$ and so by hypothesis there is a functional relation between $g_1,g_2,\ldots,g_{n-1},$ that is between $u_1,u_2,\ldots,u_{n-1}$ into which u_n may enter, because u_n may occur in set of equation (1) as an auxiliary variable. We have therefore proved by induction that there is a relation between $u_1,u_2,\ldots,u_n$.

4.9.2 Properties of Jacobian

Lemma 1. If U and V are functions of u and v, where u and v are themselves functions of x and y, we have

$$\frac{\partial(U,V)}{\partial(x,y)}=\frac{\partial(U,V)}{\partial(u,v)}.\frac{\partial(u,v)}{\partial(x,y)}$$

Proof. Let $\quad U = f(u, v),\ V = F(u, v)$

$$u=\phi(x,y),\ v=\psi(x,y)$$

Then $\quad \frac{\partial U}{\partial x}=\frac{\partial U}{\partial u}.\frac{\partial u}{\partial x}+\frac{\partial U}{\partial v}.\frac{\partial v}{\partial x}$

$$\frac{\partial U}{\partial y} = \frac{\partial U}{\partial u}.\frac{\partial u}{\partial y} + \frac{\partial U}{\partial v}.\frac{\partial v}{\partial y}$$

$$\frac{\partial V}{\partial x} = \frac{\partial V}{\partial u}.\frac{\partial u}{\partial x} + \frac{\partial V}{\partial v}.\frac{\partial v}{\partial x}$$

$$\frac{\partial V}{\partial y} = \frac{\partial V}{\partial u}.\frac{\partial u}{\partial y} + \frac{\partial V}{\partial v}.\frac{\partial v}{\partial y}$$

and

$$\frac{\partial(U,V)}{\partial(u,v)}.\frac{\partial(u,v)}{\partial(x,y)} = \begin{vmatrix} \frac{\partial U}{\partial u} & \frac{\partial U}{\partial v} \\ \frac{\partial V}{\partial u} & \frac{\partial V}{\partial u} \end{vmatrix} \times \begin{vmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{vmatrix}$$

$$= \begin{vmatrix} \frac{\partial U}{\partial u}.\frac{\partial u}{\partial x} + \frac{\partial U}{\partial v}.\frac{\partial v}{\partial x} & \frac{\partial U}{\partial u}.\frac{\partial u}{\partial y} + \frac{\partial U}{\partial v}.\frac{\partial v}{\partial y} \\ \frac{\partial V}{\partial u}.\frac{\partial u}{\partial x} + \frac{\partial V}{\partial v}.\frac{\partial v}{\partial x} & \frac{\partial V}{\partial u}.\frac{\partial u}{\partial y} + \frac{\partial V}{\partial v}.\frac{\partial v}{\partial y} \end{vmatrix}$$

$$= \begin{vmatrix} \frac{\partial U}{\partial x} & \frac{\partial U}{\partial y} \\ \frac{\partial V}{\partial x} & \frac{\partial V}{\partial y} \end{vmatrix} = \frac{\partial(U,V)}{\partial(x,y)}$$

The same method of proof applies if there are several functions and the same number of variables.

Lemma 2. If J is the Jacobian of system u, v with regard to x, y and J′ the Jacobian of x, y with regard to u, v, then J J′ = 1.

Proof. Let u = f(x, y) and v = F(x, y), and suppose that these are solved for x and y giving

$$x = \phi(u,v) \text{ and } y = \psi(u,v),$$

we then have differentiating u = f(x, y) w.r.t u and v; v = F(x, y) w.r.t u and v

$$\left.\begin{aligned} 1 &= \frac{\partial u}{\partial x}.\frac{\partial x}{\partial u} + \frac{\partial u}{\partial y}.\frac{\partial y}{\partial u} \\ 0 &= \frac{\partial u}{\partial x}.\frac{\partial x}{\partial v} + \frac{\partial u}{\partial y}.\frac{\partial y}{\partial v} \end{aligned}\right\} \text{ obtained from } u = f(x, y)$$

$$\left.\begin{aligned} 0 &= \frac{\partial v}{\partial x}.\frac{\partial x}{\partial u} + \frac{\partial v}{\partial y}.\frac{\partial y}{\partial u} \\ 1 &= \frac{\partial v}{\partial x}.\frac{\partial x}{\partial v} + \frac{\partial v}{\partial y}.\frac{\partial y}{\partial v} \end{aligned}\right\} \text{ obtained from } v = F(x, y).$$

Also $$J\,J' = \begin{vmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{vmatrix} \times \begin{vmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{vmatrix}$$

$$= \begin{vmatrix} \frac{\partial u}{\partial x}.\frac{\partial x}{\partial u} + \frac{\partial u}{\partial y}.\frac{\partial y}{\partial u} & \frac{\partial u}{\partial x}.\frac{\partial x}{\partial v} + \frac{\partial u}{\partial y}.\frac{\partial y}{\partial v} \\ \frac{\partial v}{\partial x}.\frac{\partial x}{\partial u} + \frac{\partial v}{\partial y}.\frac{\partial y}{\partial u} & \frac{\partial v}{\partial x}.\frac{\partial x}{\partial v} + \frac{\partial v}{\partial y}.\frac{\partial y}{\partial v} \end{vmatrix}$$

$$= \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = 1$$

Example 1. If $u = x + 2y + z, v = x - 2y + 3z$

$$w = 2xy - xz + 4yz - 2z^2,$$

prove that $\frac{\partial(u,v,w)}{\partial(x,y,z)} = 0$, and find a relation between u, v, w.

Solution. We have

$$\frac{\partial(u,v,w)}{\partial(x,y,z)} = \begin{vmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} & \frac{\partial u}{\partial z} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} & \frac{\partial v}{\partial z} \\ \frac{\partial w}{\partial x} & \frac{\partial w}{\partial y} & \frac{\partial w}{\partial z} \end{vmatrix}$$

$$= \begin{vmatrix} 1 & 2 & 1 \\ 1 & -2 & 3 \\ 2y - z & 2x + 4z & -x + 4y - 4z \end{vmatrix}$$

$$= \begin{vmatrix} 1 & 0 & 0 \\ 1 & -4 & 2 \\ 2y - z & 2x + 6z - 4y & -x + 2y - 3z \end{vmatrix}$$

Performing $c_2 \to c_2 - 2c_1$ and $c_3 \to c_3 - c_1$

$$= \begin{vmatrix} -4 & 2 \\ 2x + 6z - 4y & -x + 2y - 3z \end{vmatrix} = \begin{vmatrix} 0 & 2 \\ 0 & -x + 2y - 3z \end{vmatrix}$$

$= 0.$

Performing $c_1 \to c_1 + 2c_2$

Hence a relation between u, v and w exists.

Now,

$$\begin{aligned} u+v &= 2x+4z \\ u-v &= 4y-2z \\ w &= x(2y-z)+2z(2y-z) \\ &= (x+2z)(2y-z) \end{aligned}$$

$\Rightarrow$ $$4w = (u+v)(u-v)$$

$\Rightarrow$ $$4w = u^2 - v^2$$

which is the required relation.

Example 2. Find the condition that the expression $px+qy+rz, p'x+q'y+r'z$ are connected with the expression $ax^2+by^2+cz^2+2fyz+2gzx+2hxy,$ by a functional relation.

Solution. Let

$$\begin{aligned} u &= px+qy+rz \\ v &= p'x+q'y+r'z \\ w &= ax^2+by^2+cz^2+2fyz+2gzx+2hxy \end{aligned}$$

We know that the required condition is

$$\frac{\partial(u,v,w)}{\partial(x,y,z)} = 0.$$

Therefore

$$\begin{vmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} & \frac{\partial u}{\partial z} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} & \frac{\partial v}{\partial z} \\ \frac{\partial w}{\partial x} & \frac{\partial w}{\partial y} & \frac{\partial w}{\partial z} \end{vmatrix} = 0.$$

But

$$\frac{\partial u}{\partial x} = p, \frac{\partial u}{\partial y} = q, \frac{\partial u}{\partial z} = r$$

$$\frac{\partial v}{\partial x} = p', \frac{\partial v}{\partial y} = q', \frac{\partial v}{\partial z} = r'.$$

$$\frac{\partial w}{\partial x} = 2ax + 2hy + 2gz$$

$$\frac{\partial w}{\partial y} = 2hx + 2by + 2fz$$

$$\frac{\partial w}{\partial z} = 2gx + 2fy + 2cz$$

Therefore

$$\begin{vmatrix} p & q & r \\ p' & q' & r' \\ 2ax + 2hy + 2gz & 2hx + 2by + 2fz & 2gx + 2fy + 2cz \end{vmatrix} = 0$$

$$\Rightarrow \quad \begin{vmatrix} p & q & r \\ p' & q' & r' \\ a & h & g \end{vmatrix} = 0, \quad \begin{vmatrix} p & q & r \\ p' & q' & r' \\ h & b & f \end{vmatrix} = 0, \quad \begin{vmatrix} p & q & r \\ p' & q' & r' \\ g & f & c \end{vmatrix} = 0$$

which is the required condition.

Example 3. Prove that if f(0) = 0, $f'(x) = \frac{1}{1+x^2}$, then

$$f(x) + f(y) = f\left(\frac{x+y}{1-xy}\right).$$

Solution. Suppose that

$$u = f(x) + f(y)$$

$$v = \frac{x+y}{1-xy}$$

Now $$J(u, v) = \begin{vmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{vmatrix}$$

$$= \begin{vmatrix} \frac{1}{1+x^2} & \frac{1}{1+y^2} \\ \frac{1+y^2}{(1-xy)^2} & \frac{1+x^2}{(1-xy)^2} \end{vmatrix} = 0$$

Therefore u and v are connected by a functional relation

Let $u = \phi(v)$, that is,

$$f(x) + f(y) = \phi\left(\frac{x+y}{1-xy}\right)$$

Putting y = 0, we get

$$f(x) + f(0) = \phi(x)$$

$$\Rightarrow \quad f(x) + 0 = \phi(x) \text{ because } f(0) = 0$$

Hence $$f(x)+f(y)=f\left(\frac{x+y}{1-xy}\right).$$

Example 4. The roots of the equation in λ

$$(\lambda-x)^3+(\lambda-y)^3+(\lambda-z)^3=0$$

are u, v, w. Prove that $\dfrac{\partial(u,v,w)}{\partial(x,y,z)}=-2\dfrac{(y-z)(z-x)(x-y)}{(v-w)(w-u)(u-v)}$.

Solution. Here u, v, w are the roots of the equation

$$\lambda^3-(x+y+z)\lambda^2+(x^2+y^2+z^2)\lambda-\frac{1}{3}(x^3+y^3+z^3)=0$$

Let $$x+y+z=\xi,\ \ x^2+y^2+z^2=\eta,\ \ \frac{1}{3}(x^3+y^3+z^3)=\zeta \tag{1}$$

and then $$u+v+w=\xi,\ \ vw+wu+uv=\eta,\ \ uvw=\zeta \tag{2}$$

Then from (1),

$$\frac{\partial(\xi,\eta,\zeta)}{\partial(x,y,z)}=\begin{vmatrix}1 & 1 & 1\\ 2x & 2y & 2z\\ x^2 & y^2 & z^2\end{vmatrix}=2(y-z)(z-x)(x-y) \tag{3}$$

Again, from (2), we have

$$\frac{\partial(\xi,\eta,\zeta)}{\partial(u,v,w)}=\begin{vmatrix}1 & 1 & 1\\ v+w & w+u & u+v\\ vw & wu & uv\end{vmatrix}=-(v-w)(w-u)(u-v) \tag{4}$$

Then from (3) and (4)

$$\frac{\partial(u,v,w)}{\partial(x,y,z)}=\frac{\partial(u,v,w)}{\partial(\xi,\eta,\zeta)}.\frac{\partial(\xi,\eta,\zeta)}{\partial(x,y,z)}=-2\frac{(y-z)(z-x)(x-y)}{(v-w)(w-u)(u-v)}$$

Example 5. If α,β,γ are the roots of the equation $\dfrac{x}{a+k}+\dfrac{y}{b+k}+\dfrac{z}{c+k}=1$ in k,

then

$$\frac{\partial(x,y,z)}{\partial(\alpha,\beta,\gamma)}=-\frac{(\alpha-\beta)(\beta-\gamma)(\gamma-\alpha)}{(a-b)(b-c)(c-a)}.$$

Solution. The equation in k is

$$k^3+k^2(a+b+c-x-y-z)+k[ab+bc+ca-x(b+c)-y(c+a)-z(a+b)] \tag{1}$$
$$+abc-bcx-cay-abz=0.$$

Now α, β, γ are the roots of this equation. Therefore

$$\alpha + \beta + \gamma = -(a + b + c) + x + y + z$$
$$\alpha\beta + \beta\gamma + \gamma\alpha = ab + bc + ca - x(b + c) - y(c + a) - z(a + b)$$

and

$$\alpha\beta\gamma = -abc + bcx + cay + abz$$

Then, we have

$$1 = \frac{\partial x}{\partial \alpha} + \frac{\partial y}{\partial \alpha} + \frac{\partial z}{\partial \alpha}$$
$$1 = \frac{\partial x}{\partial \beta} + \frac{\partial y}{\partial \beta} + \frac{\partial z}{\partial \beta}$$
$$1 = \frac{\partial x}{\partial \gamma} + \frac{\partial y}{\partial \gamma} + \frac{\partial z}{\partial \gamma}$$
$$\beta + \gamma = -(b + c)\frac{\partial x}{\partial \alpha} - (c + a)\frac{\partial y}{\partial \alpha} - (a + b)\frac{\partial z}{\partial \alpha}$$
$$\gamma + \alpha = -(b + c)\frac{\partial x}{\partial \beta} - (c + a)\frac{\partial y}{\partial \beta} - (a + b)\frac{\partial z}{\partial \beta}$$
$$\alpha + \beta = -(b + c)\frac{\partial x}{\partial \gamma} - (c + a)\frac{\partial y}{\partial \gamma} - (a + b)\frac{\partial z}{\partial \gamma}$$
$$\beta\gamma = bc\frac{\partial x}{\partial \alpha} + ca\frac{\partial y}{\partial \alpha} + ab\frac{\partial z}{\partial \alpha}$$
$$\gamma\alpha = bc\frac{\partial x}{\partial \beta} + ca\frac{\partial y}{\partial \beta} + ab\frac{\partial z}{\partial \beta}$$
$$\alpha\beta = bc\frac{\partial x}{\partial \gamma} + ca\frac{\partial y}{\partial \gamma} + ab\frac{\partial z}{\partial \gamma}$$

Now,
$$\begin{vmatrix} \frac{\partial x}{\partial \alpha} & \frac{\partial y}{\partial \alpha} & \frac{\partial z}{\partial \alpha} \\ \frac{\partial x}{\partial \beta} & \frac{\partial y}{\partial \beta} & \frac{\partial z}{\partial \beta} \\ \frac{\partial x}{\partial \gamma} & \frac{\partial y}{\partial \gamma} & \frac{\partial z}{\partial \gamma} \end{vmatrix} \begin{vmatrix} 1 & 1 & 1 \\ -(b + c) & -(c + a) & -(a + b) \\ bc & ca & ab \end{vmatrix}$$

$$= \begin{vmatrix} 1 & 1 & 1 \\ \beta + \gamma & \gamma + \alpha & \alpha + \beta \\ \beta\gamma & \gamma\alpha & \alpha\beta \end{vmatrix}$$

Hence

$$\frac{\partial(x, y, z)}{\partial(\alpha, \beta, \gamma)}(b - c)(c - a)(a - b) = -(\alpha - \beta)(\beta - \gamma)(\gamma - \alpha)$$

$$\Rightarrow \quad \frac{\partial(x, y, z)}{\partial(\alpha, \beta, \gamma)} = -\frac{(\alpha - \beta)(\beta - \gamma)(\gamma - \alpha)}{(b - c)(c - a)(a - b)}$$

Second Method. After the equation (1),

let $a+b+c-(x+y+z)=\xi$

$ab+bc+ca-x(b+c)-y(c+a)-z(a+b)=\eta$

$$abc-bcx-cay-abz=\zeta \tag{2}$$

$$\alpha+\beta+\gamma=-\xi,\ \ \alpha\beta+\beta\gamma+\gamma\alpha=\eta,\ \alpha\beta\gamma=-\zeta. \tag{3}$$

Then

$$\frac{\partial(\xi,\eta,\zeta)}{\partial(x,y,z)}=\begin{vmatrix}-1 & -1 & -1\\ -(b+c) & -(c+a) & -(a+b)\\ -bc & -ca & -ab\end{vmatrix}=(a-b)(b-c)(c-a)\,.$$

and

$$\frac{\partial(\xi,\eta,\zeta)}{\partial(\alpha,\beta,\gamma)}=\begin{vmatrix}-1 & -1 & -1\\ \beta+\gamma & \gamma+\alpha & \alpha+\beta\\ -\beta\gamma & -\gamma\alpha & -\alpha\beta\end{vmatrix}$$

$$=\ -(\alpha-\beta)(\beta-\gamma)(\gamma-\alpha)$$

Therefore $$\frac{\partial(x,y,z)}{\partial(\alpha,\beta,\zeta)}=\frac{\partial(x,y,z)}{\partial(\xi,\eta,\zeta)}.\frac{\partial(\xi,\eta,\zeta)}{\partial(\alpha,\beta,\gamma)}=-\frac{(\alpha-\beta)(\beta-\gamma)(\gamma-\alpha)}{(a-b)(b-c)(c-a)}.$$

Example 6. Prove that the three functions U, V, W are connected by an identical functional relation if

$$U = x + y - z,\ V = x - y + z,\ W = x^2 +y^2+z^2-2yz$$

and find the functional relation.

Solution. Here

$$\frac{\partial(U,V,W)}{\partial(x,y,z)}=\begin{vmatrix}\frac{\partial U}{\partial x} & \frac{\partial U}{\partial y} & \frac{\partial U}{\partial z}\\ \frac{\partial V}{\partial x} & \frac{\partial V}{\partial y} & \frac{\partial V}{\partial z}\\ \frac{\partial W}{\partial x} & \frac{\partial W}{\partial y} & \frac{\partial W}{\partial z}\end{vmatrix}$$

$$=\begin{vmatrix}1 & 1 & -1\\ 1 & -1 & 1\\ 2x & 2(y-z) & 2(z-y)\end{vmatrix}$$

Performing $c_3 \to c_3 + c_2$

$$=\begin{vmatrix}1 & 1 & 0\\ 1 & -1 & 0\\ 2x & 2(y-z) & 0\end{vmatrix}=0$$

Hence there exists some functional relation between U, V and W.

Moreover,

$$U + V = 2x$$

$$U - V = 2(y - z)$$

$$(U+V)^2 + (U-V)^2 = 4(x^2 + y^2 + z^2 - 2yz)$$

$$= 4W$$

which is the required functional relation.

Example 7. Let V be a function of the two variables x and y. Transform the expression

$$\frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2}$$

by the formulae of plane polar transformation

$$x = r\cos\theta, \quad y = r\sin\theta .$$

Solution. We are given a function V which is function of x and y and therefore it is a function of r and θ. From $x = r\cos\theta, \quad y = r\sin\theta$, we have

$$r = \sqrt{x^2 + y^2}, \quad \theta = \tan^{-1} y/x.$$

Now

$$\frac{\partial V}{\partial x} = \frac{\partial V}{\partial r}.\frac{\partial r}{\partial x} + \frac{\partial V}{\partial \theta}.\frac{\partial \theta}{\partial x}$$

$$= \cos\theta \frac{\partial V}{\partial r} - \frac{\sin\theta}{r}\frac{\partial V}{\partial \theta} \qquad \left(\because \frac{\partial r}{\partial x} = \cos\theta, \quad \frac{\partial \theta}{\partial x} = -\frac{\sin\theta}{r}\right)$$

And

$$\frac{\partial V}{\partial y} = \frac{\partial V}{\partial r}.\frac{\partial r}{\partial y} + \frac{\partial V}{\partial \theta}.\frac{\partial \theta}{\partial y}$$

$$= \sin\theta \frac{\partial V}{\partial r} + \frac{\cos\theta}{r}\frac{\partial V}{\partial \theta} \qquad \left(\because \frac{\partial r}{\partial y} = \sin\theta, \quad \frac{\partial \theta}{\partial y} = \frac{\cos\theta}{r}\right)$$

Therefore

$$\frac{\partial}{\partial x} = \left(\cos\theta \frac{\partial}{\partial r} - \frac{\sin\theta}{r}\frac{\partial}{\partial \theta}\right)$$

$$\frac{\partial}{\partial y} = \left(\sin\theta \frac{\partial}{\partial r} + \frac{\cos\theta}{r}\frac{\partial}{\partial \theta}\right)$$

Hence

$$\frac{\partial^2 V}{\partial x^2} = \left(\cos\theta \frac{\partial}{\partial r} - \frac{\sin\theta}{r}\frac{\partial}{\partial \theta}\right)\left(\cos\theta \frac{\partial V}{\partial r} - \frac{\sin\theta}{r}\frac{\partial V}{\partial \theta}\right)$$

$$= \cos\theta \frac{\partial}{\partial r}\left(\cos\theta \frac{\partial V}{\partial r} - \frac{\sin\theta}{r}\frac{\partial V}{\partial \theta}\right) - \frac{\sin\theta}{r}\frac{\partial}{\partial \theta}\left(\cos\theta \frac{\partial V}{\partial r} - \frac{\sin\theta}{r}\frac{\partial V}{\partial \theta}\right)$$

$$= \cos\theta\left(\cos\theta\frac{\partial^2 V}{\partial r^2} + \frac{\sin\theta}{r^2}\frac{\partial V}{\partial \theta} - \frac{\sin\theta}{r}\frac{\partial^2 V}{\partial r\partial\theta}\right)$$

$$-\frac{\sin\theta}{r}\left(\cos\theta\frac{\partial^2 V}{\partial\theta\partial r} - \sin\theta\frac{\partial V}{\partial r} - \frac{\cos\theta}{r}\frac{\partial V}{\partial\theta} - \frac{\sin\theta}{r}\frac{\partial^2 V}{\partial\theta^2}\right)$$

$$= \cos^2\theta\frac{\partial^2 V}{\partial r^2} - 2\frac{\sin\theta\cos\theta}{r}\frac{\partial^2 V}{\partial r\partial\theta} + \frac{\sin^2\theta}{r^2}\frac{\partial^2 V}{\partial\theta^2}$$

$$+\frac{\sin^2\theta}{r}\frac{\partial V}{\partial r} + \frac{2\sin\theta\cos\theta}{r^2}\frac{\partial V}{\partial\theta} \qquad (1)$$

and
$$\frac{\partial^2 V}{\partial y^2} = \left(\sin\theta\frac{\partial}{\partial r} + \frac{\cos\theta}{r}\frac{\partial}{\partial\theta}\right)\left(\sin\theta\frac{\partial V}{\partial r} + \frac{\cos\theta}{r}\frac{\partial V}{\partial\theta}\right)$$

$$= \sin\theta\frac{\partial}{\partial r}\left(\sin\theta\frac{\partial V}{\partial r} + \frac{\cos\theta}{r}\frac{\partial V}{\partial\theta}\right) + \frac{\cos\theta}{r}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial V}{\partial r} + \frac{\cos\theta}{r}\frac{\partial V}{\partial\theta}\right)$$

$$= \sin\theta\left(\sin\theta\frac{\partial^2 V}{\partial r^2} - \frac{\cos\theta}{r^2}\frac{\partial V}{\partial\theta} + \frac{\cos\theta}{r}\frac{\partial^2 V}{\partial r\partial\theta}\right)$$

$$+\frac{\cos\theta}{r}\left(\sin\theta\frac{\partial^2 V}{\partial\theta\partial r} + \cos\theta\frac{\partial V}{\partial r} - \frac{\sin\theta}{r}\frac{\partial V}{\partial\theta} + \frac{\cos\theta}{r}\frac{\partial^2 V}{\partial\theta^2}\right),$$

$$= \sin^2\theta\frac{\partial^2 V}{\partial r^2} + \frac{\sin\theta\cos\theta}{r}\frac{\partial^2 V}{\partial r\partial\theta} - \frac{\cos\theta\sin\theta}{r^2}\frac{\partial V}{\partial\theta}$$

$$+\frac{\cos\theta\sin\theta}{r}\frac{\partial^2 V}{\partial\theta\partial r} + \frac{\cos^2\theta}{r^2}\frac{\partial^2 V}{\partial\theta^2} + \frac{\cos^2\theta}{r}\frac{\partial V}{\partial r}$$

$$-\frac{\sin\theta\cos\theta}{r^2}\frac{\partial V}{\partial\theta} \qquad (2)$$

Adding (1) and (2), we obtain

$$\frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} = \frac{\partial^2 V}{\partial r^2} + \frac{1}{r}.\frac{\partial V}{\partial r} + \frac{1}{r^2}\frac{\partial^2 V}{\partial\theta^2}$$

which is the required result.

Example 8. Transform the expression

$$\left(x\frac{\partial Z}{\partial x} + y\frac{\partial Z}{\partial y}\right)^2 + (a^2 - x^2 - y^2)\left\{\left(\frac{\partial Z}{\partial x}\right)^2 + \left(\frac{\partial Z}{\partial y}\right)^2\right\}$$

by the substitution $x = r\cos\theta$, $y = r\sin\theta$.

Solution. If V is a function of x, y, then

$$\frac{\partial V}{\partial r} = \frac{\partial V}{\partial x}\frac{\partial x}{\partial r} + \frac{\partial V}{\partial y}\frac{\partial y}{\partial r} = \frac{x}{r}\frac{\partial V}{\partial x} + \frac{y}{r}\frac{\partial V}{\partial y}$$

$$\Rightarrow \qquad r\frac{\partial V}{\partial r} = x\frac{\partial V}{\partial x} + y\frac{\partial V}{\partial y} = \left(x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y}\right)V$$

$$\Rightarrow \qquad r\frac{\partial}{\partial r} = x\frac{\partial}{\partial x} + y\frac{\partial}{\partial y}$$

Similarly $\qquad \dfrac{\partial}{\partial \theta} = x\dfrac{\partial}{\partial y} - y\dfrac{\partial}{\partial x}$

Now $\qquad \dfrac{\partial Z}{\partial x} = \dfrac{\partial Z}{\partial r}.\dfrac{\partial r}{\partial x} + \dfrac{\partial Z}{\partial \theta}.\dfrac{\partial \theta}{\partial x} = \cos\theta\dfrac{\partial Z}{\partial r} - \dfrac{\sin\theta}{r}\dfrac{\partial Z}{\partial \theta} \qquad (1)$

$$\frac{\partial Z}{\partial y} = \sin\theta\frac{\partial Z}{\partial r} + \frac{\cos\theta}{r}\frac{\partial Z}{\partial \theta} \qquad (2)$$

Therefore $\qquad \left(\dfrac{\partial Z}{\partial x}\right)^2 + \left(\dfrac{\partial Z}{\partial y}\right)^2 = \left(\dfrac{\partial Z}{\partial r}\right)^2 + \dfrac{1}{r^2}\left(\dfrac{\partial Z}{\partial \theta}\right)^2$

and the given expression is equal to

$$\left(r\frac{\partial Z}{\partial r}\right)^2 + (a^2 - r^2)\left[\left(\frac{\partial Z}{\partial r}\right)^2 + \frac{1}{r^2}\left(\frac{\partial Z}{\partial \theta}\right)^2\right]$$

$$= a^2\left(\frac{\partial Z}{\partial r}\right)^2 + \left(\frac{a^2}{r^2} - 1\right)\left(\frac{\partial Z}{\partial \theta}\right)^2.$$

Example 9. If $x = r\cos\theta$, $y = r\sin\theta$, prove that

$$(x^2 - y^2)\left(\frac{\partial^2 u}{\partial x^2} - \frac{\partial^2 u}{\partial y^2}\right) + 4xy\frac{\partial^2 u}{\partial x \partial y} = r^2\frac{\partial^2 u}{\partial r^2} - r\frac{\partial u}{\partial r} - \frac{\partial^2 u}{\partial \theta^2}$$

where u is any twice differentiable function of x and y.

Solution. We have

$$\frac{\partial u}{\partial r} = \frac{\partial u}{\partial x}.\frac{\partial x}{\partial r} + \frac{\partial u}{\partial y}.\frac{\partial y}{\partial r}$$

$$= \cos\theta\frac{\partial u}{\partial x} + \sin\theta\frac{\partial u}{\partial y} = \frac{x}{r}\frac{\partial u}{\partial x} + \frac{y}{r}\frac{\partial u}{\partial y}$$

$$\Rightarrow \qquad r\frac{\partial u}{\partial r} = x\frac{\partial u}{\partial x} + y\frac{\partial u}{\partial y} \qquad (1)$$

Therefore $\qquad r\dfrac{\partial}{\partial r}\left(r\dfrac{\partial u}{\partial r}\right) = \left(x\dfrac{\partial}{\partial x} + y\dfrac{\partial}{\partial y}\right)\left(x\dfrac{\partial u}{\partial x} + y\dfrac{\partial u}{\partial y}\right)$

$$= x\frac{\partial}{\partial x}\left(x\frac{\partial u}{\partial x} + y\frac{\partial u}{\partial y}\right) + y\frac{\partial}{\partial y}\left(x\frac{\partial u}{\partial x} + y\frac{\partial u}{\partial y}\right)$$

$$= x^2\frac{\partial^2 u}{\partial x^2} + xy\frac{\partial^2 u}{\partial x\partial y} + xy\frac{\partial^2 u}{\partial y\partial x} + y^2\frac{\partial^2 u}{\partial y^2} + x\frac{\partial u}{\partial x} + y\frac{\partial u}{\partial y}$$

Therefore

$$\therefore \qquad r^2\frac{\partial^2 u}{\partial r^2} + r\frac{\partial u}{\partial r} = x^2\frac{\partial^2 u}{\partial x^2} + 2xy\frac{\partial^2 u}{\partial x\partial y} + y^2\frac{\partial^2 u}{\partial y^2} + x\frac{\partial u}{\partial x} + y\frac{\partial u}{\partial y} \qquad (2)$$

$$r^2\frac{\partial^2 u}{\partial r^2} = x^2\frac{\partial^2 u}{\partial x^2} + 2xy\frac{\partial^2 u}{\partial x\partial y} + y^2\frac{\partial^2 u}{\partial y^2} \qquad \text{(using (1))}$$

Again, $$\frac{\partial u}{\partial \theta} = \frac{\partial u}{\partial x}.\frac{\partial x}{\partial \theta} + \frac{\partial u}{\partial y}.\frac{\partial y}{\partial \theta}$$

$$= x\frac{\partial u}{\partial y} - y\frac{\partial u}{\partial x}$$

Therefore $$\frac{\partial^2 u}{\partial \theta^2} = \left(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}\right)\left(x\frac{\partial u}{\partial y} - y\frac{\partial u}{\partial x}\right)$$

$$= x\frac{\partial}{\partial y}\left(x\frac{\partial u}{\partial y} - y\frac{\partial u}{\partial x}\right) - y\frac{\partial u}{\partial x}\left(x\frac{\partial u}{\partial y} - y\frac{\partial u}{\partial x}\right)$$

$$= x^2\frac{\partial^2 u}{\partial y^2} - 2xy\frac{\partial^2 u}{\partial y\partial x} + y^2\frac{\partial^2 u}{\partial y^2} - x\frac{\partial u}{\partial x} - y\frac{\partial u}{\partial y} \qquad (3)$$

From (1), (2) and (3), we get the required result.

4.10 References

1. Walter Rudin, Principles of Mathematical Analysis (3rd edition) McGraw-Hill, Kogakusha, 1976, International Student Edition.
2. T. M. Apostol, Mathematical Analysis, Narosa Publishing House, New Delhi, 1974.
3. H.L. Royden, Real Analysis, Macmillan Pub. Co., Inc. 4th Edition, New York, 1993.
4. G. De Barra, Measure Theory and Integration, Wiley Eastern Limited, 1981.
5. R.R. Goldberg, Methods of Real Analysis, Oxford & IBH Pub. Co. Pvt. Ltd, 1976.
6. R. G. Bartle, The Elements of Real Analysis, Wiley International Edition, 2011.
7. S.C. Malik and Savita Arora, Mathematical Analysis, New Age International Limited, New Delhi, 2012.